Journal of Experimental Psychology:
Learning, Memory, and Cognition
2000, Vol. 26, No. 5, 1141–1159

0278-7393/00/$5.00 DOI: 10.1037//0278-7393.26.5.1141

Similarity and Inhibition in Long-Term Memory: Evidence for a Two-Factor Theory

Michael C. Anderson and Collin Green
University of Oregon

Kathleen C. McCulloch
New York University

Recalling a past experience often requires the suppression of related memories that compete with the retrieval target, causing memory impairment known as retrieval-induced forgetting. Two experiments examined how retrieval-induced forgetting varies with the similarity of the competitor and the target item (target–competitor similarity) and with the similarity between the competitors themselves (competitor–competitor similarity). According to the pattern-suppression model (M. C. Anderson & B. A. Spellman, 1995), high target–competitor similarity should reduce impairment, whereas high competitor–competitor similarity should increase it. Both predictions were supported: Encoding target–competitor similarities not only eliminated retrieval-induced forgetting but also reversed it, whereas encoding competitor–competitor similarities increased impairment. The differing effects of target–competitor and competitor–competitor similarity may resolve conflicting results concerning the effects of similarity on inhibition.

Bringing a memory to consciousness often involves the discrimination of the event we want from other similar memory traces. Retrieving the memory for where we parked today, for instance, may require that we overcome the repeated recollection of where we parked yesterday. Instances of retrieval interference such as this pervade cognition, whether one is attempting to recall events, facts, locations, speech sounds, or even motor responses (see Anderson & Neely, 1996, for a review; see also Crowder, 1976; Postman, 1971; Postman & Underwood, 1973; Roediger & Neely, 1982). Although we often resolve these confusions quickly, the mechanisms that allow us to overcome interference from competing memory traces have enduring consequences for the state of memory.

A variety of findings have shown that overcoming interference from a related memory while recalling a retrieval target can render that related memory less accessible. Evidence that retrieval processes impair interfering memories has been found in studies of episodic recall (Anderson, Bjork, & Bjork, 1994; Anderson & Spellman, 1995; Bauml, 1998; Roediger & Schmidt, 1980; A. D. Smith, 1971) and semantic generation (Blaxton & Neely, 1983; Brown, 1981; Dagenbach, Carr, & Barnhardt, 1990), as well as in studies using verbal (Anderson et al., 1994; Macrae & MacLeod, 1999; R. E. Smith & Hunt, in press), visuospatial (Ciranni & Shimamura, 1999), and complex eyewitness event stimuli as materials (Koutstaal, Schacter, Johnson, & Galluccio, 1999; Shaw, Bjork, & Handal, 1995). The existence of this form of impairment has been taken by many as evidence for inhibitory processes that suppress competing memories (Anderson et al., 1994; Anderson & Spellman, 1995; Bauml, 1996, 1997, 1998; Blaxton & Neely, 1983; Dagenbach et al., 1990; Macrae & MacLeod, 1999). Because it is initiated by the retrieval process itself and because its effects have been shown to last at least 20 min (Anderson et al., 1994; Anderson & Spellman, 1995), this phenomenon is often referred to as *retrieval-induced forgetting* (Anderson et al., 1994).

The tendency for retrieval to impair the later recall of related memories supports the notion that inhibitory processes help to discriminate targets from similar competing traces in memory (see, e.g., Anderson & Spellman, 1995; Dagenbach et al., 1990). However, if inhibitory processes function to discriminate targets from competitors, one might expect to observe more retrieval-induced forgetting for memories that are highly similar to a retrieval target than for memories that are less similar to it. More impairment should be observed for highly similar items because activating a similar competitor in memory should trigger inhibitory processes to a greater extent than activating a less similar one. Although much of the data on retrieval-induced forgetting are broadly consistent with this expectation, recent studies that have looked at the effect of interitem similarity on retrieval-induced forgetting have yielded inconsistent findings. In some studies, interitem similarity has been found to increase retrieval-induced forgetting (R. E. Smith & Hunt, in press), whereas in others it has been found to reduce or eliminate it altogether (Hartinger & Bauml, 1999). In still other studies, using an output interference procedure, intercategory similarity had little effect on the amount of output interference (Roediger & Schmidt, 1980).

In the present article, we examine the relation between similarity and retrieval-induced forgetting and attempt to reconcile these contradictory findings. First, we review recent work using the

Michael C. Anderson and Collin Green, Department of Psychology, University of Oregon; Kathleen C. McCulloch, Department of Psychology, New York University.

This research was supported in part by startup funds from the University of Oregon.

We thank Bobbie Spellman and Chad Dodson for their comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Michael C. Anderson, Department of Psychology, University of Oregon, 1227 University of Oregon, Eugene, Oregon 97403-1227. Electronic mail may be sent to mcanders@darkwing.uoregon.edu.

retrieval-practice paradigm of Anderson et al. (1994), that has been used to study the effects of similarity on inhibition. We then suggest that the two conclusions that have been drawn from these studies—that similarity works to increase and decrease impairment—are both correct. We believe that these contradictory findings illustrate the influence of two different dimensions of similarity that affect observed impairment in opposite ways: *target–competitor similarity* and *competitor–competitor similarity*. We then report two experiments that conceptually replicate the major findings of R. E. Smith and Hunt (in press) and Hartinger and Bauml (1999) and that tie those findings to these two types of similarity. In this way, we hope to reconcile these findings with each other and with classical data on the role of similarity in interference.

Similarity and Retrieval-Induced Forgetting

In research on retrieval-induced forgetting, it is often found that retrieval impairs items that are semantically similar to the retrieved target more than items that are unrelated to it (Anderson et al., 1994; Anderson et al., in press; Anderson & McCulloch, 1999; Anderson & Spellman, 1995; R. E. Smith & Hunt, in press; see also Blaxton & Neely, 1983; Dagenbach et al., 1990; but see Roediger & Schmidt, 1980). This pattern can be illustrated with the retrieval-practice paradigm of Anderson et al. (1994). In the standard version of this paradigm, there are four phases: the study, retrieval-practice, distractor, and final test phases. In the study phase, subjects encode six members from each of eight taxonomic categories (e.g., *Fruits, Trees*), which are presented in category–exemplar paired-associates format (e.g., *Fruit Orange*). Next, they are given an opportunity to perform retrieval practice on some of the items they studied. Subjects usually practice recalling half of the members from half of the categories by means of recurring category-plus-stem cued recall tests on these items (e.g., *Fruit Or—* for *Orange*). After a 20-min distractor phase, sujects are given a category-cued recall test for all of the studied categories. Final recall is measured on three types of items: (a) practiced items from practiced categories (hereinafter called *practiced items;* e.g., *Fruit Orange*), (b) unpracticed items from practiced categories (unpracticed competitor items; e.g., *Fruit Banana*), and (c) unpracticed items from unpracticed study categories (*baseline items;* e.g., *Drinks Scotch*). As might be expected, retrieval practice typically facilitates the later recall of practiced items (e.g., *Fruit Orange*). More interesting, however, is the finding that retrieval practice impairs recall of the unpracticed competitors relative to recall for baseline items. Importantly, because impairment is measured against the recall of within-subjects baseline items, these findings show that retrieval-induced forgetting can be category specific, consistent with a dependency of impairment on similarity.

If the presence of similar memories trigger inhibitory processes, one might expect that as competitors are made more similar to a retrieval target, retrieval-induced forgetting for those competitors should increase. Recently, two studies have looked at this hypothesis by adapting the retrieval-practice procedure so that the degree of within-category similarity might be varied. R. E. Smith & Hunt, (in press) altered the study phase to encourage the encoding of similarities or differences between exemplars of a category. For the similarity encoding group, subjects viewed all six exemplars of the category at once and were asked to find a way that the item at the top of the list was similar to all of the remaining items. Shared features were then generated in turn for the other five exemplars. After encoding the categories in this way, subjects went through the remaining phases of the retrieval-practice procedure. The difference encoding group followed the same steps but was asked instead to find one feature that made the top item different from all of the remaining items. R. E. Smith and Hunt's results supported their hypothesis: Encoding differences abolished retrieval-induced forgetting, but encoding similarities yielded robust impairment. R. E. Smith and Hunt suggested that distinctive encoding reduced the interference that ordinarily triggers inhibition. In support of this view, they cited classical studies showing that retroactive interference increases with the degree of interlist similarity (McGeoch & McDonald, 1931; McGeoch & McGeoch, 1936; Shuell, 1968).

Although R. E. Smith and Hunt's (in press) findings appear to support the discrimination view, a different conclusion emerges from two experiments reported by Hartinger and Bauml (1999). Like R. E. Smith and Hunt, these authors manipulated the degree of similarity between items in a category (e.g., *Fruits*). However, instead of manipulating encoding instructions, Hartinger and Bauml used the standard intentional learning procedure of the retrieval-practice procedure and varied whether or not unpracticed competitors (e.g., *Fruit Lemon*) were drawn from the same subcategory (e.g., *Citrus*) as a to-be-practiced item (e.g., *Fruit Orange*) or a different one (e.g., *Fruit Cherry*). After studying the items, subjects engaged in retrieval practice and then, after a 2-min retention interval, were given a final category-plus-stem cued recall test (e.g., *Fruit O—*). In their first experiment, Hartinger and Bauml found retrieval-induced forgetting regardless of whether retrieval-practice targets and their competitors shared a subcategory, but sharing a subcategory did reduce impairment nonsignificantly. In a second experiment, the item similarity manipulation was strengthened by highlighting the crucial similarities during the study phase: Each category–exemplar pair was listed with the item's subcategory name (e.g., *Fruit Citrus Orange*). Consistent with their first experiment, retrieval-induced forgetting was eliminated when unpracticed items shared a subcategory with practiced items, but not when they were drawn from different subcategories, in contrast to what might be expected on Smith and Hunt's data and the discrimination view.

Given the apparent consistency of R. E. Smith and Hunt's (in press) results with findings from the classical interference era, one might suspect that the Hartinger and Bauml (1999) data are not representative. There are good reasons to reject this idea, however. First, although some findings from the interference era support R. E. Smith and Hunt's conclusions, others do not and are more in line with Hartinger and Bauml's findings. For instance, many studies using the paired-associates method have found that when response words in a second list of paired associates (e.g., *Dog–Couch*) are similar to the response words in a first list (e.g., *Dog–Chair*), retroactive interference can be eliminated (Dallett, 1962; Kanungo, 1967; Morgan & Underwood, 1950; Osgood, 1946, 1948; Postman, 1964; Young, 1955), even with extensive training on the interpolated list (e.g., Barnes & Underwood, 1959). During the interference era, this contradictory pattern of similarity effects was noted in a classic article by Osgood (1949), who referred to a "similarity paradox" in studies of interference. (We return to Osgood's treatment of the paradox in the General Discussion.) Second, and perhaps more directly relevant, Anderson and McCulloch (1999) recently showed that instructing subjects to rehearse and interrelate category exemplars during the study phase eliminated retrieval-induced forgetting. Although Anderson and McCulloch did not intend to investigate similarity, it seems likely that their subjects interrelated exemplars by finding properties the items had in common, suggesting that similarity reduces retrieval-induced forgetting, as suggested by Hartinger and Bauml.

The findings of R. E. Smith and Hunt (in press) and Hartinger and Bauml (in press) thus pose a contradiction that is not easily resolved by appeal to related findings in the classical interference literature, which is itself complex. The resolution of this question must instead come from a reconsideration of the effects of similarity on the degree of retrieval-induced forgetting. That concern is the subject of our next section.

Similarity and Retrieval-Induced Forgetting Reconsidered

According to the discrimination view of retrieval-induced forgetting, activating a similar competitor should trigger inhibitory processes to a greater extent than activating a less similar one. Thus, if retrieval-induced forgetting reflects the amount of suppression exerted on an item, similar competitors should be more impaired. One problem with this argument, however, is that it does not consider the effects of a successful retrieval practice on those features that a competitor has in common with the practiced item.

The effects of shared features on retrieval-induced forgetting is best illustrated in terms of the distributed model suggested by Anderson and Spellman (1995), depicted in Figure 1. According to this model, items are represented as sets of features that are encoded when an item gets studied. The item's representation will include those features to which a subject attends, and items that are similar to one another will tend to share feature units. Retrieving a target involves the activation of all and only those features included in that item's pattern. When attempts to retrieve a target activate similar patterns, interference occurs, necessitating the suppression of features not shared with the target item. The persisting effect of this suppression is thought to contribute to retrieval-induced forgetting.

A different effect occurs for those features that a competitor has in common with the retrieval target. After a target is retrieved, all of its feature units are strengthened, which causes the facilitation typically exhibited by items that have been practiced (Allen, Mahler, & Estes, 1969; Bjork, 1975; Carrier & Pashler, 1992; Gardiner, Craik, & Bleasdale, 1973; Hogan & Kintsch, 1971). Because of this strengthening, however, features of the competing item that overlap with the target are facilitated, as illustrated by the darkened circles in Figure 1a. This facilitation, together with the suppression of a competitor's distinctive features, will determine the amount of retrieval-induced forgetting that is observed. Thus, when the effects of retrieval on overlapping features are considered, a different prediction about the effects of similarity emerges: Making items very similar should reduce, not increase, impairment. Consider Figure 1b, in which a competitor and a target differ by only a few features. Because these patterns are so similar, selectively retrieving the target should be difficult, and the suppression exerted on the competitor's distinctive features should be greater than in the example depicted in Figure 1a. However, because retrieving the target strengthens most of the competitor's features, the net effect of this similarity will be to reduce or perhaps even reverse retrieval-induced forgetting. Thus, although some amount of similarity between a target and a competitor seems necessary for competition to ensue (and for inhibition to be needed), the Anderson and Spellman (1995) model, in line with the results of Hartinger and Bauml (1999), predicts less impairment as competitors become highly similar to retrieval targets. Indeed, Hartinger and Bauml explained their findings by appeal to the preceding implication of the Anderson and Spellman model.

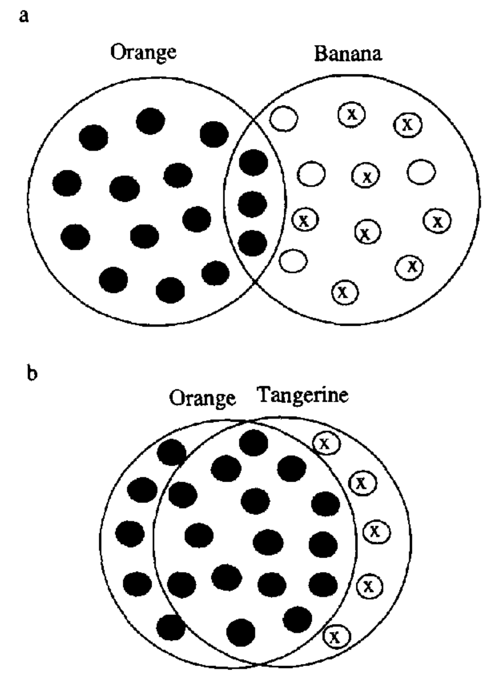

Figure 1. Low (a) and high (b) degrees of target–competitor similarity, as conceived in the Anderson and Spellman (1995) model. Larger circles represent individual category exemplars, smaller circles represent semantic features, darkened circles represent the strengthening of a feature as a result of retrieval practice, and Xs denote that a feature has been suppressed. *Orange* is the retrieval-practice target; *Banana* (for Figure 1a) and *Tangerine* (for Figure 1b) are the unpracticed competitors. Note that the high degree of target–competitor similarity in Figure 1b (represented by a high number of overlapping features) causes a substantial proportion of *Tangerine*'s features to be strengthened by retrieval practice rather than be suppressed.

The foregoing account cannot explain the discrepancy between the R. E. Smith and Hunt (in press) and Hartinger and Bauml (1999) findings, however. The clear reduction in retrieval-induced forgetting observed by R. E. Smith and Hunt with distinctive encoding suggests that another aspect of similarity must be considered. The answer may lie in the distinction between *target–competitor similarity*[1] (i.e., similarity between a retrieval-practice

[1] Throughout this article, we use the term *target* to refer to an item that has or will receive retrieval practice and the term *competitor* to refer to the nonpracticed items from the same category as the retrieval-practice targets. This terminology is emphasized here to prevent confusion due to the fact that on the final recall test, the competitors themselves become the retrieval targets (the item to be retrieved).

target and an unpracticed competitor), the focus of discussion so far, and what we call *competitor–competitor similarity.* When a retrieval-practice target has more than one competitor, the competitors themselves can vary in how similar they are to one another. This fact is illustrated in Figures 2a and 2b, which depict low and high competitor–competitor similarity, respectively. As can be seen in Figure 2, the effect of competitor–competitor similarity on retrieval-induced forgetting should be opposite to that of target–competitor similarity. When competitors overlap a lot (Figure 2b), suppression of the same number of features should have a greater negative effect than when competitors do not overlap (Figure 2a). This should occur because when a feature is suppressed, it will remain suppressed throughout the experiment. Thus, any other patterns that also share that feature will be impaired as a result. In effect, suppressing a feature in one pattern can do "double duty" by impairing all other competitors that share it. Thus, when an item has a similar competitor, a greater proportion of its features will be suppressed, yielding more retrieval-induced forgetting.

We suspect that R. E. Smith and Hunt's (in press), encoding instructions to make each exemplar different from every other one in the category may have decreased both target–competitor and competitor–competitor similarity. If so, distinctive encoding may not have reduced retrieval-induced forgetting by making competitors more discriminable from targets but rather by diminishing competitor–competitor similarity. If this is true, then if we manipulate target–competitor and competitor–competitor similarity separately, we should be able to show that the former kind of similarity reduces retrieval-induced forgetting, as observed by Hartinger and Bauml, whereas the latter kind increases it, as observed by R. E. Smith and Hunt.

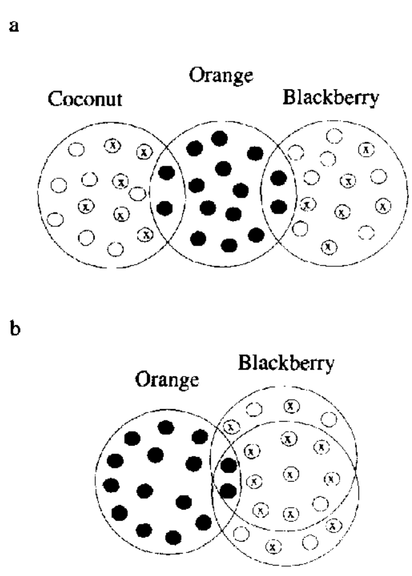

Figure 2. Low (a) and high (b) degrees of competitor–competitor similarity, as conceived in the Anderson and Spellman (1995) model. Larger circles represent individual category exemplars, smaller circles represent semantic features, darkened circles represent the strengthening of a feature as a result of retrieval practice, and Xs denote that a feature has been suppressed. Note the greater number of features overlapping between *Blackberry* and *Raspberry* in Figure 2b than between *Blackberry* and *Coconut* in Figure 2a. In both figures, the same number of features has been suppressed (12 features); in Figure 2b, suppressing the same number of features results in a greater proportion of each competitor's features being suppressed than in Figure 2a.

The Present Experiments

In this article, we report two experiments that examine the effects of similarity on retrieval-induced forgetting. We adapted the retrieval-practice procedure of Anderson et al. (1994) to manipulate target–competitor and competitor–competitor similarity separately. To see whether variations in target–competitor and competitor–competitor similarity affect retrieval-induced forgetting differently, we modified the study phase of the retrieval-practice procedure. The new encoding procedure required two passes through the materials. In the first pass, subjects were presented with exemplars and asked to judge how good each example was as a member of its category. This exemplar-goodness-judgment task encouraged subjects to encode each exemplar in relation to its category name, increasing the likelihood that items would become associated to their categories (and thus compete later on during the retrieval-practice task). The incidental nature of this encoding task also reduced the chances that subjects would integrate exemplars with one another, a strategy typical of subjects in the standard intentional learning procedure (Anderson & McCulloch, 1999). If subjects' tendency to interrelate exemplars was left unchecked, they might encode similarities between unspecified pairs of items, contaminating our effort to control which patterns of similarity they encoded in the second study pass.

In the second pass through the materials, we manipulated target–competitor and competitor–competitor similarity. Following R. E. Smith and Hunt (in press), we varied interitem similarity by giving subjects an encoding task in which they looked for similarities or differences between exemplars. However, instead of asking subjects to make each exemplar similar (or different) from every other, we drew their attention to target–competitor similarities for some categories (i.e., *Heart–Radish* in Figure 3a) but competitor–competitor similarities for others (i.e., *Tomato–Radish* in Figure 3b). On each page of the second encoding booklet was printed a category along with an exemplar pair (e.g., *Red Tomato Radish*). For categories in the target–competitor condition, the pairs were composed of one exemplar from the half of the category that was to receive retrieval practice in the following phase (i.e., the target) and one of its unpracticed competitors; for categories in the competitor–competitor condition, both items in a pair were either to-be-practiced items or were to-be-competitors. Baseline categories for each of these conditions were constructed similarly, except that no items received practice (see the introduction to Experiment 1 for a more complete description of these phases, as well as a graphical summary in Figure 4). To increase either kind of similarity, subjects were asked to find as many similarities as they could between the items in a pair (Experiment 1); to decrease similarity, they were asked to find unique properties of each item (Experiment 2).

A second aim of these experiments was to see whether variations in retrieval-induced forgetting arising from our manipulations of similarity are cue independent. In prior work, Anderson and Spellman (1995) found that retrieval-induced forgetting gen-

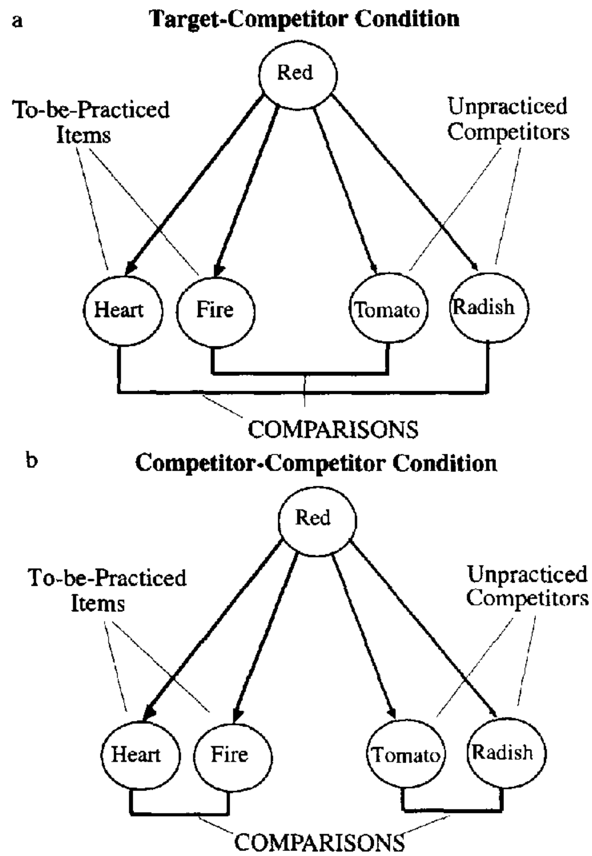

Figure 3. A schematic representation of the target–competitor condition (a) and the competitor–competitor condition (b). To simplify the presentation, four exemplars are shown for a single category—two to-be-practiced items (left sides of Figures 3a and 3b) and two unpracticed competitors (right sides of Figures 3a and 3b). In the second encoding phase, for categories in the target–competitor condition (Figure 3a), subjects are presented with pairs composed of one target and one competitor, as denoted by the solid line at the bottom linking items from each set. Similarly, for categories in the competitor–competitor condition (Figure 3b), subjects are presented with pairs composed of two to-be-practiced targets (left side) and pairs composed of two unpracticed competitors (right side), as denoted by the solid lines at the bottom that link exemplars. Note that every exemplar participates in a comparison and that the same number of comparisons are encoded in each case—only the pattern of comparison (target–competitor vs. competitor–competitor) varies.

eralized to retrieval cues other than the ones used to do retrieval practice. In particular, they showed that retrieval practice on some members of a category (e.g., *Red Heart,* through cues like *Red He*—) impaired recall of other exemplars (e.g., *Tomato, Radish*), regardless of whether the final recall of those items was tested with the same category used to do retrieval practice (e.g., the cue *Red* for the item *Tomato*) or a different intralist category cue (e.g., the cue *Food* for *Radish*). This finding is significant because it rules out noninhibitory sources of interference that may contribute to retrieval-induced forgetting (for a review of these noninhibitory alternatives, see Anderson & Bjork, 1994). For instance, on the final test, when subjects try to recall *Tomato* given the cue *Red,* they may fail to recall *Tomato* either because it was suppressed in the practice phase or because its recall was blocked by stronger items during the final test. *Tomato* might be blocked if using the cue *Red* led to the persistent intrusion of items like *Red–Heart,* which became powerfully associated to that cue in the earlier practice phase. By testing *Tomato* with *Food* instead of *Red, Heart* will not block *Tomato* because *Food* is not associated with that item, making it unlikely for *Heart* to intrude in response to that cue. Thus, the independent probe method provides a measurement of the degree of inhibition that is uncontaminated by other noninhibitory influences and can establish whether variations in impairment are cue independent.

To implement the independent probe method in Experiments 1 and 2 in the present study, we designed categories containing eight exemplars. For each category, four of its eight items also fell under an implicit category. For example, the category *Red* was designed to have eight members, four of which were also members of the implicit category *Food* (e.g., *Cherry, Radish, Tomato,* and *Apple;* hereinafter called *cross-categorizable items*) and four of which were not (e.g., *Brick, Fire, Sunburn,* and *Heart;* hereinafter called *regular items*). After encoding categories like this, subjects did retrieval practice on the practiced target items (which were always the regular items, e.g., *Red–Heart*) by means of category-stem cues (e.g., *Red He*—). On the final test, instead of giving subjects the studied category name (*Red*) as a cue, we gave them the name of the implicit category (e.g., *Food*), along with the first letter for each of the unpracticed competitors or baseline items (which, in this study, were always the cross-categorizable items) they had encoded. Subjects were told that these category names were new but that many of the previously judged items could fall under them and that they should try to recall any item they had seen that fit the cues. Thus, the implicit category name served as an independent probe for the unpracticed competitors (and cross-categorizable items in baseline categories) because it was associated with them but not with the practiced items. If retrieval practice suppresses unpracticed competitors, these items should be recalled more poorly than baseline items, even though they are tested with a novel extralist category cue.

Experiment 1

In Experiment 1, we divided the encoding phase of the retrieval-practice procedure into two parts—an initial encoding phase (Phase I of Figure 4) and a similarity encoding phase (Phase II of Figure 4). In the initial encoding phase, subjects were presented with word triplets composed of a category name and two exemplars (e.g., *Red Brick Tomato*) and were asked to decide which item was a better example of the category. This exemplar-goodness judgment directed subjects, attention to the relation between each item and the category, encouraging the encoding of category–exemplar associations. We chose to have subjects make these judgments on pairs of exemplars (relative goodness judgments) so that we could carefully direct subjects' nearly uncontrollable tendency to make comparative goodness ratings (e.g., "This example is better than the one I saw before"). If subjects' comparative tendencies are not carefully directed, their spontaneous comparisons with previous exemplars might lead them to incidentally encode similarities between unspecified pairs of items. Such uncontrolled comparisons, much like intentional integration, would undo our efforts to precisely control the particular patterns

PHASES OF THE PROCEDURE

I. INITIAL ENCODING PHASE (note: underlined cross-categorizable items all belong to another single implicit category)

Target-Competitor Categories		**Competitor-Competitor Categories**	
To-be Practiced	*Baseline*	*To-be Practiced*	*Baseline*
Red brick tomato	Wood fence bench	Fly blimp frisbee	Soft grass skin
Red apple heart	Wood desk log	Fly glider kite	Soft hair pillow
Red fire cherry	Wood mast cabinet	Fly eagle bat	Soft satin flannel
Red radish sunburn	Wood stool crate	Fly wasp owl	Soft velvet cotton
Sharp thorn lance	Loud siren jackhammer	Religious rosary priest	Rocks granite slate
Sharp bayonet needle	Loud drill boom	Religious altar bible	Rocks pumice shale
Sharp scalpel sword	Loud yell compressor	Religious mosque synagogue	Rocks diamond amethyst
Sharp dagger tack	Loud lawnmower traffic	Religious convent temple	Rocks opal emerald

Task: Circle which item is the better exemplar of the category.
Aim: To get subjects to encode each exemplar in association with its category name.

II. SIMILARITY ENCODING PHASE All of the pairs reappear in a second booklet, with a rating scale. The aim is to get subjects to encode similarities between particular pairs of items. For instance:

Red tomato brick 0 1 2 (a target-competitor item)
Fly bat eagle 0 1 2 (a competitor-competitor item)

Task: Find as many "similarities between the two examples as you can, aside from their membership in the sated category. Circle 0, 1, or 2.

III. RETRIEVAL PRACTICE PHASE. All nonunderlined regular items from to-be-practiced categories are given retrieval practice.

Practiced Target-Competitor Categories		*Practiced Competitor-Competitor Categories*	
Red Brick	Sharp Thorn	Fly Blimp	Religious Rosary
Red Heart	Sharp Needle	Fly Frisbee	Religious Priest
Red Fire	Sharp Scalpel	Fly Glider	Religious Altar
Red Sunburn	Sharp Tack	Fly Kite	Religious Bible

Task: Recall the item you saw that fits the cues provided. (e.g., Red Br____)

20 Minute Retention Interval

IV. EXTRALIST CUED RECALL FINAL TEST PHASE

Task: Recall the previously seen item that can fall under the new category that begins with the letter provided. All underlined items are tested.

Target-Competitor Condition		**Competitor-Competitor Condition**	
Unpracticed Competitors	*Baseline*	*Unpracticed Competitor*	*Baseline*
Food (T, A, C, R)	Furniture (B, D, C, S)	Animal (E, W, B, O)	Cloth (V, F, S, C)
Weapon (L, B, S, D)	Tools (J, D, C, L)	Building (M, T, S, C)	Gems (D, E, A, O)

Figure 4. A graphical summary of the phases of Experiment 1, with representative materials in each condition. Phases I and II are both encoding phases, with the former designed to foster category-exemplar associations and the latter to implement our critical similarity encoding. Target–competitor categories present subjects with pairs composed of one regular and one cross-categorizable item (underlined). Competitor–competitor categories present subjects with pairs of regular items or of cross-categorizable items. During Phase III, subjects engaged in retrieval practice on the regular (nonunderlined) items from two target–competitor and two competitor–competitor categories. After a 20-min retention interval, the critical cross-categorizable (underlined) items were tested with an extralist category cue.

of similarity (i.e., target–competitor or competitor–competitor similarity) that subjects were to encode in the subsequent similarity encoding phase. For these reasons, we had subjects make relative goodness decisions for the very same pairs of items for which they were to later encode similarities (Phase II of Figure 4), making any unintended similarity encoding for those item pairs work in concert with our later similarity encoding manipulation.

After encoding category–exemplar associations, we implemented the critical similarity encoding manipulation in a second pass through the materials (using the same triplets as were used in the initial encoding phase). This similarity encoding phase is illustrated in the Figure 4 (Phase II). For the sake of illustration, the initial encoding portion of Figure 4 (Phase I) lists all of the categories and exemplars that a single subject might have seen (except filler categories), with four categories assigned to each of the target–competitor (left side of figure) and competitor–competitor (right side) similarity encoding conditions. (Note, however, that although these items are classified into these encoding conditions in the initial encoding section of this figure, the instructions to encode similarities took place in Phase II.) As can be seen in this figure, subjects thought of similarities for exemplar pairs composed of a to-be-practiced target item (i.e., regular items, which are not underlined in the figure) and a competitor (i.e., cross-categorizable items, which are underlined in the figure) in the target–competitor condition; for the remaining categories, subjects thought of similarities between pairs of to-be-practiced targets (nonunderlined items) and pairs of competitors (underlined items).

After completing the similarity encoding phase, subjects performed retrieval practice on four of the eight categories (Phase III of Figure 4)—two target–competitor categories and two competitor–competitor categories—according to the procedure of Anderson et al. (1994). For each practiced category, subjects recalled the four regular (nonunderlined) items three times each, with the remaining four cross-categorizable (underlined) items serving in the unpracticed competitor condition.

In the final test phase (Phase IV of Figure 4), each trial cued subjects with the implicit category name for a given study category, together with the first letter of a cross-categorizable item. Exemplars were tested one at a time, in blocks of four from the same implicit category. Because the regular items from each of the original categories were not members of implicit categories, they were not tested. The first letter of each exemplar was given because pilot studies showed that cuing with an extralist category by itself led subjects to supplement their recall by covertly cuing themselves with the original category names. Because we intended the extralist category to be an independent probe for the unpracticed competitors, we tried to minimize this strategy. By providing a letter stem, we hoped to give subjects enough information so that self-cuing would be less necessary. As an additional control, we limited each test trial to 5 s, assuming that a brisk pace would discourage elaborate self-cuing strategies. To verify that our controls were effective, we added a final questionnaire by which we could measure cuing.

If retrieval-induced forgetting is found, we should observe that retrieval practice impairs final recall for unpracticed competitors relative to items from baseline categories. If retrieval-induced forgetting occurs, it would extend the cue-independent impairment observed by Anderson and Spellman (1995) to our new extralist cuing procedure and also demonstrate that our new incidental encoding method by itself does not eliminate impairment. Crucially, if high target–competitor similarity reduces retrieval-induced forgetting, as suggested by the Anderson and Spellman model and by Hartinger and Bauml's (1999) findings, impairment should be less in the target–competitor condition than in the competitor–competitor condition. Indeed, retrieval-induced forgetting may be eliminated or even reversed in the target–competitor condition.

Method

Subjects

Forty-eight University of Oregon undergraduates participated to fulfill a course requirement.

Design

Two factors, pattern-of-comparison and retrieval-practice status, were manipulated within subjects. Pattern-of-comparison had two levels: target–competitor similarity and competitor–competitor similarity. For both types of similarity, subjects were presented with pairs of exemplars and asked to find similarities between them. For categories in the target–competitor condition, subjects found similarities between pairs composed of a regular item and a cross-categorizable item. In the competitor–competitor condition, subjects were asked to find similarities between items both drawn from the regular or the cross-categorizable set.

The retrieval-practice status of an item was manipulated within subjects. Exemplars either (a) were unpracticed but members of a practiced category (unpracticed competitors) or (b) were unpracticed and members of an unpracticed study category (baseline items). Items that received retrieval practice (practiced items) were not tested on the final recall test because they were not members of implicit categories, and so they were not included in the design. The dependent measure was the proportion of items correctly recalled in each condition on an extralist category-plus-stem cued recall test.

In addition to these manipulations, we examined how subjects' strategies in the test phase might modulate inhibition effects. To achieve this, we divided subjects into groups on the basis of their responses to the postexperimental questionnaire. For each question, we divided subjects into three groups (low, medium, and high) by (a) sorting subjects within each counterbalancing group by their score for the question of interest and (b) assigning the bottom, middle, and top thirds to the low, moderate, and high groups, respectively.

Materials and Procedure

Figure 4 displays the materials and procedure of Experiment 1 and is used throughout this section for illustration.

Category construction. Ten categories, 2 of which were fillers, were constructed, Phase I of Figure 4 (initial encoding phase) presents these eight categories along with the eight exemplars constructed for each. Four of the eight items selected for each category were chosen so that they could fall under a second (implicit) category (see underlined items in the figure). For example, for the category red things, we chose four examples that could also be categorized as foods (e.g., *Tomato, Radish, Cherry,* and *Apple*), called *cross-categorizable items.* The remaining four examples were designed not to be members of that second category (e.g., not foods—*Brick, Sunburn, Heart,* and *Fire*) and are called *regular items*.[2] For each explicit category, the implicit category that went with it was chosen so as to reduce its association with the explicit category (e.g., *Food* and *Red* are not tightly associated).

In addition to the above constraints, care was taken to ensure that no member of a category was a member of any other explicit or implicit category in the experiment. Within each category, a priori interassociations between exemplars were avoided whenever possible and exemplars were constrained to be single words. An effort was made to ensure that every studied item began with a distinct two-letter stem, ensuring that stems were unique for retrieval practice. Some overlapping two-letter stems were allowed, but only in the cross-categorizable item sets, which never received retrieval practice (and only between category overlaps were allowed). Stem difficulty was controlled by using stems with high versatility (i.e., number of words in Kučera & Francis, 1967, with that stem- [Solso & Juel, 1980]; versatility, $M = 237$). In addition, items within the cross-categorizable set of each explicit category were constrained to begin with a distinct first letter.

Initial encoding phase. In the initial encoding phase, subjects were randomly assigned to one of two learning booklets, differing only in which categories were assigned to the target–competitor and competitor–competitor conditions. Each of these booklets contained 80 items, 8 items from eight experimental and two filler categories. Each page of the booklet displayed a category with two exemplars (e.g., *Red Heart Tomato*). As can be seen in the top of Figure 4 for categories in the competitor–competitor condition (right half of Figure 4), the two exemplars in a given pair were drawn from only one subset—either the cross-categorizable item set (e.g., *Fly Eagle Bat*—see underlined pairs) or the regular item set (e.g., *Fly Glider Kite*—see nonunderlined pairs). For categories in the target–competitor condition (left half of Figure 4, top), the two exemplars were drawn from the two different category subsets (e.g., *Red Brick Tomato*). For each of the pages, subjects were given 3 s to circle the exemplar that they thought was the better member of the category.[3] Subjects were told that after 3 s, a beep would signal them to move on to the next pair in the list and that this would continue until the list was completed.

For target–competitor pairs, the cross-categorizable item appeared as the first example for one half of the items in a category and as the second example for the other half (see left side, top, Figure 4). The exemplars were presented in block-randomized format, with 1 pair from each category appearing in each block of trials. This yielded four blocks of 10 pairs each, equating each category and condition for serial position within the learning order. To control for primacy and recency effects, the first and last two items in the list were fillers.

Similarity encoding phase. As indicated in the similarity encoding Phase II of Figure 4, subjects received the pairs that were presented in the initial encoding phase a second time. The second booklet presented the exemplar pairs in the same block-randomized order. During this second phase, however, subjects were told that they would be given 5 s for each triplet to think of as many "similarities" between the two examples as they could. It was explained that similarities were characteristics or parts that the two examples had in common, other than their membership in the listed category. Subjects indicated their response by circling a number on a 3-point scale (0 = *no similarities*, 1 = *1 similarity*, 2 = *greater than 1 similarity*) printed next to each triplet, indicating how many similarities they generated (see Figure 4). Subjects were told that after 6 s, a beep would signal them to move to the next pair. Aside from this change in page format and procedure, the only other difference from the initial encoding phase was that the order in which exemplars were presented in a given pair (e.g., *Red Tomato Brick*) was inverted relative to the order in the initial encoding phase (e.g., *Red Brick Tomato*).

Retrieval-practice phase. In the retrieval-practice phase, subjects were told that each page of the booklet they received contained a category with the first two letters of an exemplar presented earlier (e.g., *Red He*—), which they were to recall and write down within 10 s. Subjects were warned that items would be tested several times and that they should try to write the correct item each time.

The 72 test pages in each practice booklet were ordered according to several constraints. To control for primary and recency effects, the first and last few pages tested filler items. Critical items were tested three times, with tests ordered by an expanding schedule; on average, 3.3 items intervened between the first and second practice test and 6.6 items intervened between the second and third. No two items from a category were tested adjacently, and the mean test position of categories was matched ($M = 31.1$). When possible, we kept sequences of tests from repeating by adding tests of fillers.

As can be seen from the example given in Figure 4 (see the Phase III table), subjects did retrieval practice on the regular items from two target–competitor categories and two competitor–competitor categories. The remaining two target–competitor and competitor–competitor categories served as baselines for those conditions. To ensure that every item appeared in every condition,

[2] The regular items were also different from the cross-categorizable items in that the former were not members of their own implicit category. Thus, cross-categorizable items were more similar to one another than were regular items. This difference in the properties of these sets was difficult to avoid, given the variety of constraints that we imposed on the construction of the categories and their exemplars. However, because every category participated in every condition, this difference between these item sets is held constant across all of our conditions. It thus seems unlikely that this difference could have contributed to the findings observed in Experiments 1 and 2. Nevertheless, it remains possible that our conclusions are specific to the use of categories with heterogeneous practice sets. However, post hoc analyses indicated that all of crucial findings reported in Experiments 1 and 2 as present in the overall recall data were also present in those categories with the most homogeneous sets of practiced items.

[3] Although using an incidental exemplar rating task in this initial encoding phase is likely to get rid of intentional integration between exemplars, it is possible that incidental integration might still take place. Incidental integration might occur if subjects rate the goodness of a given exemplar by comparison to how good they thought other previously rated exemplars of that category were. The act of comparing two category members to construct a new rating might have the same functional effect as intentionally linking them together. Because it is difficult to prevent subjects from making comparative ratings, we chose instead to encourage them but to control which item they compared each exemplar with. This is why in the initial exemplar rating phase subjects see two exemplars (the same two that they will later compare explicitly) and are asked to judge which one is the better exemplar, instead of rating each exemplar individually on a Likert scale. By doing this, we prevent subjects from making comparisons with other exemplars, which could ruin our attempts to separate target–competitor and competitor–competitor similarity.

we counterbalanced which categories were practiced. One half of the subjects practiced the categories depicted in Phase III of Figure 4, and the other half practiced the remaining four categories. This counterbalancing required two retrieval practice booklets, each containing 16 exemplars (4 exemplars from each of four categories) ordered according to the constraints described previously. Subjects were randomly assigned to one of these two booklets at the outset of the retrieval-practice procedure. After the practice phase, subjects did a causal reasoning task for the 20-min retention interval. The reasoning booklet included no words used in the main experiment.

Test phase. For the final test, subjects were told that each page of the test booklet contained a new category they had not seen before, along with the first letter of a previously encountered item that could fall under it (see Figure 4, bottom, for examples). For each of these test pages, subjects were given 5 s to use the new category and the letter stem as cues to recall an earlier studied item. After 5 s, a beep signaled subjects to proceed to the next page. Subjects were given only 5 s per example to reduce their ability to use complex search strategies, such as scanning through the previously studied category names to generate additional cues.

In the test booklet, exemplars from each of the new category names were tested on separate pages, with a single-letter stem printed next to the implicit category name (e.g., *Food C—*). Because only half of the items in each originally studied category (e.g., *Red*) were designed to be cross-categorizable items (e.g., members of *Food*), this final test only assessed memory for four of the eight items in each of the original categories (see Figure 4). The cross-categorizable items were tested consecutively in a block of four test trials, after which a test of a different category proceeded.

To control output interference across practiced and baseline categories, the mean test position of the practiced and baseline sets was matched. The order of particular categories was also counterbalanced so that the position of every category was equated across subjects. On the basis of these constraints, two test books were made.

Questionnaire. After the experiment, subjects completed a questionnaire in which they were asked about strategies and experiences during some phases of the experiment. The questionnaire asked people about two strategies, one that may have arisen in the retrieval-practice phase and one that may have arisen in the final test phase. The retrieval-practice phase question asked subjects to estimate how much time they spent engaged in "extra retrieval practice." It was explained that extra retrieval practice meant any additional rehearsal on examples other than the one being tested on a given page after the target item had been recalled. Subjects' tendency to adopt this strategy was measured by asking subjects to rate each category for how often they did extra retrieval practice.

The question concerning the final test asked people whether they tried to scan back through the originally encoded categories to help them think of responses to the new implicit category names. We measured this "covert cuing" by having subjects rate each implicit category separately for the degree of covert cuing they did. Subjects made their responses on a 5-point scale (1 = *none of the time,* 3 = *some of the time,* 5 = *all of the time).* Subjects were given the following description of covert cuing, from which they were asked to base their ratings:

> When I saw each category on the final test, I mentally scanned through the earlier category names that I had seen to get ideas about what to put down. When members of those earlier categories came to mind, I decided whether to write them down, depending on whether they fit.

Each question was read aloud by the experimenter while the subjects followed along. To ensure that subjects spent enough time on their answers, they were given a fixed amount of time to respond to each question. Subjects were asked to answer honestly and accurately and were told that there were no right or wrong answers to the questions.

Results and Discussion

All of the analyses were performed with study and practice counterbalancing as between-subjects variables. An alpha level of .05, two-tailed, was used for all of the statistical tests.

Retrieval Practice

The percentage of items recalled during the retrieval-practice phase did not vary across the competitor–competitor (M = 81%) and target–competitor (M = 77%) conditions ($F < 1$).

Final Recall Performance

Final recall performance is shown in Figure 5a. Overall recall performance was marginally better in the competitor–competitor condition (M = 46%) than in the target–competitor condition (42%), $F(1, 40) = 3.10$, $p = .09$, $MSE = 0.029$.

Effects of retrieval practice. The final recall of unpracticed competitors was unimpaired by retrieval practice in the overall analysis (unpracticed competitor − baseline = 43% − 45% = −2%); $F < 1$), suggesting little evidence of inhibition in this procedure. However, this overall comparison masks a striking interaction between the competitor–competitor and target–competitor conditions in the amount of inhibition suffered, $F(1, 40) = 13.90$, $MSE = 0.025$. As can be seen in Figure 5, the competitor–competitor condition showed significant retrieval-induced forgetting (unpracticed competitor − baseline = 41% − 51% = −10%), $F(1, 40) = 8.95$, $MSE = 0.026$. The finding of inhibition in the competitor–competitor condition, despite the use of a novel category cue, extends prior work establishing that retrieval-induced forgetting is cue independent (Anderson & Spellman, 1995) by showing that cue-independent impairment can also be observed with an extralist retrieval cue. It further shows that making similarity judgments in the encoding phase does not by itself eliminate inhibition. In the target–competitor condition, however, similarity judgments had a powerful effect on inhibition: Unpracticed competitors were not merely insulated from inhibition but also were reliably facilitated by retrieval practice of their category mates (unpracticed competitor − baseline = 45% − 38% = 7% facilitation), $F(1, 40) = 7.48$, $MSE = 0.016$. Thus, encouraging the encoding of similarities between the retrieval-practice targets and competitors reversed retrieval-induced forgetting, consistent with what is expected based on the Anderson and Spellman model.

Although retrieval practice appears to facilitate unpracticed competitors in the target–competitor condition, one might be con-

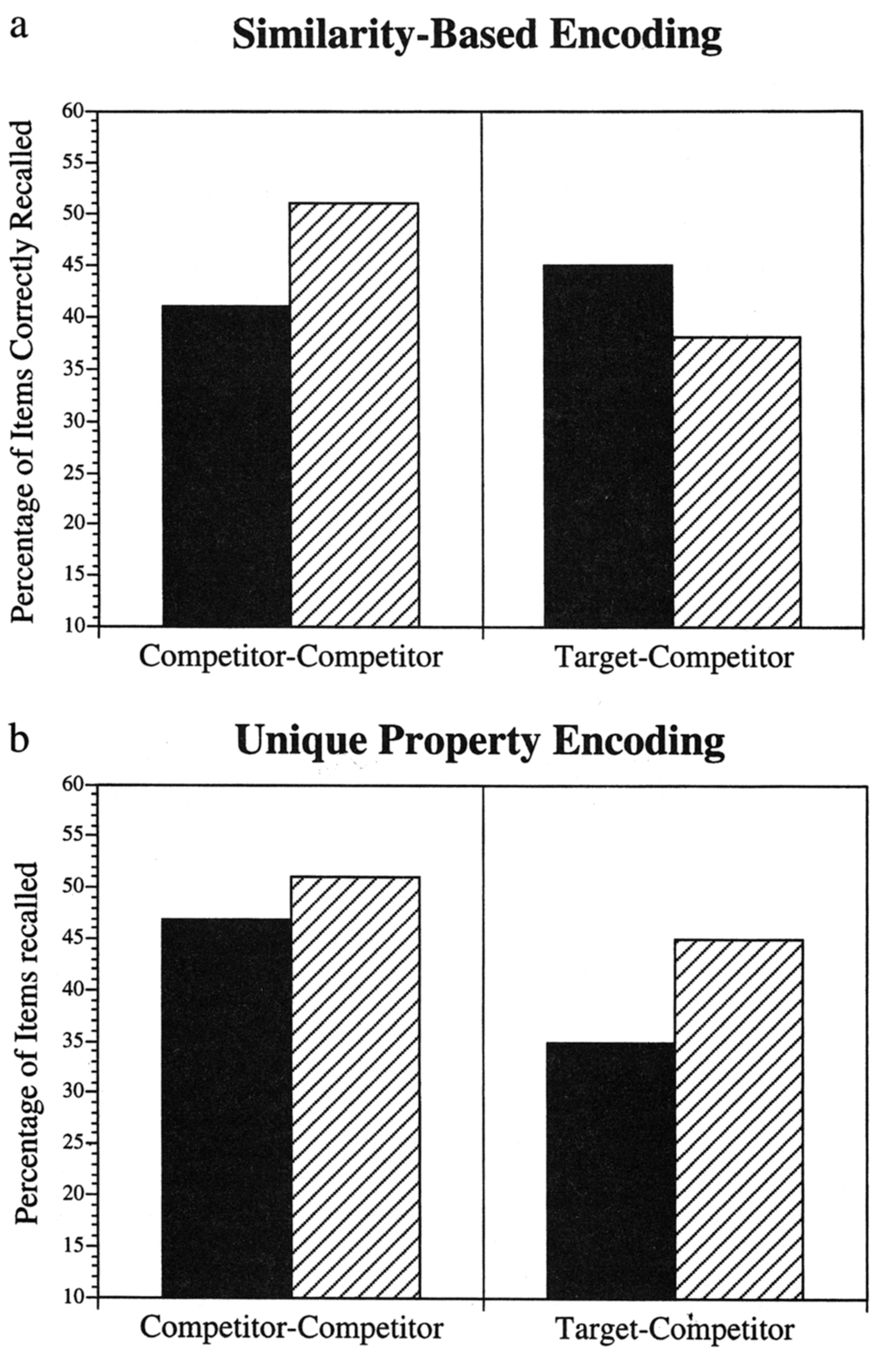

Figure 5. The effects of pattern-of-comparison on retrieval-induced forgetting in (a) Experiment 1 (similarity-based encoding) and (b) Experiment 2 (unique property encoding). The left two bars in Figures 5a and 5b depict performance on unpracticed competitor items (unpracticed members of practiced categories—black bar) and baseline items (unpracticed members of unpracticed categories—lined bar) for subjects in the competitor–competitor condition. The right two bars in each figure depict performance of unpracticed competitor and baseline items for subjects in the target–competitor condition.

cerned about the manner in which this facilitation effect appears to be generated. As can be seen in Figure 5a, the facilitation effect in this condition (right-hand side) appears to be generated more by a drop in baseline performance (38%), relative to the competitor–competitor condition (51%), than by an absolute increase in the recall of unpracticed competitor items (45% and 41% in the target–competitor and competitor–competitor conditions, respectively). One might argue that if the baseline had not dropped, normal inhibition effects would have been found. Although this concern seems reasonable at first, one must consider that the critical cross-categorizable items that are measured on our final recall test were encoded differently in the target–competitor and competitor–competitor conditions. In the competitor–competitor condition, subject's attention was deliberately drawn to similarities between the cross-categorizable items themselves (e.g., *Cherry, Radish*), a process likely to generate encodings compatible with the extralist category cue (e.g., *Food*). In the target–competitor condition, however subjects compared cross-categorizable and regular (non-cross-categorizable) items (e.g., *Sunburn, Radish*). This difference in context is likely to have discouraged (or at least to not have encouraged) the encoding of features relevant to the extralist category cue. For these reasons, overall recall should be expected to be lower in the target–competitor condition, as indeed it was. However, this lower overall recall rate in the target–competitor condition is not relevant to the critical finding. In that condition, both the unpracticed competitors and the baseline items were encoded in exactly the same way (and with exactly the same disadvantage) and were thus matched in every respect, except retrieval practice on the retrieval-practice targets. Thus, the recall advantage of unpracticed competitors over baseline items in that condition may be taken as a genuine facilitation of those items relative to their expected baseline level of performance, given the type of encoding performed.

The reversal of retrieval-induced forgetting in the target–competitor condition illustrates how important the pattern of similarity is to determining the amount of retrieval-induced forgetting. When similarities were emphasized in the competitor–competitor condition, we observed normal amounts of inhibition, consistent with R. E. Smith and Hunt's (in press) finding that similarity-based encoding failed to reduce retrieval-induced forgetting. When similarities were emphasized in the target–competitor condition, however, inhibition turned into facilitation, consistent with Hartinger and Bauml's (1999) finding that interitem similarity reduced retrieval-induced forgetting. Thus, whether similarity increases or decreases inhibition depends on the pattern of similarity that one manipulates. Nevertheless, both effects of similarity are compatible with the discrimination hypothesis proposed by R. E. Smith and Hunt, provided that the effects of retrieval practice on overlapping features are considered: Even if high target–competitor similarity decreases the discriminability of the target item and increases the need to suppress the competitor, greater retrieval-induced forgetting will not be observed if enough of that competitor's features are shared by the target and that target is strengthened by retrieval practice (i.e., Figure 1b).

A feature-based interpretation of the present inhibition and facilitation effects seems especially appropriate given our use of the independent probe method on the final test. As we argued previously, testing an unpracticed competitor (e.g., *Red Tomato*) from another retrieval cue (e.g., *Food*) that is not related to practiced items (e.g., *Heart*) allows us to measure the state of the practiced competitor independent of interference from practiced associations (e.g., *Red Heart*). That unpracticed competitors are still impaired (competitor–competitor condition) or facilitated (target–competitor condition) from an independent cue is thus consistent with some change to the state of the unpracticed competitor itself. A change in the activation of constituent features such as that proposed by the Anderson and Spellman (1995) model provides a straightforward explanation for such effects, particularly given the reversal of inhibition with similarity-based processing in the target–competitor condition.

One might argue that the facilitatory effects of retrieval practice on unpracticed competitors could still be explained by associative mechanisms if subjects used more retrieval cues than were provided on the final test. For instance, to recall members of the extralist category (e.g., *Food*), subjects may have tried to recall earlier studied category names (e.g., *Red*) to help generate candidates that could fit the new category and the letter stem. If this strategy enabled subjects to recall a practiced item (e.g., *Red Heart*), they might have then used an associative pathway from that item to recall the unpracticed competitor with which it was compared (e.g., *Tomato*). Because practiced items will be more accessible than corresponding items in baseline categories, this covert self-cuing strategy should be more useful for unpracticed competitors than for baseline items. Thus, covert cuing might compensate for inhibition of unpracticed competitors. To determine whether covert cuing is important to the present results, we analyzed the recall data in light of postexperimental questionnaire responses.

Postexperimental questionnaire analysis. Analyses based on postexperimental questionnaire responses suggest that this covert cuing strategy does not explain the target–competitor facilitation effect. First, the addition of letter-stem cues on the final recall test together with the reduction in the recall time given per item were effective in reducing covert cuing as a strategy relative to our pilot study. This finding can be seen in the lower reported self-cuing in Experiment 1 ($M = 2.68$ on a 5-point scale, with 3 labeled *some of the time*) than in pilot studies in which subjects (a) received extralist category cues without constraining letter stems and (b) received 30 s per category of recall time (7.5 s per item; $M = 3.47$), $F(1, 96) = 16.00$, $MSE = 1.07$.[4] Second, when our subject group was divided into thirds, on the basis of the reported amount of self-cuing (low third = 1.63 rating, high third = 3.57 rating), no reliable differences in facilitation or inhibition across these groups were found. If anything, inhibition and facilitation in the competitor–competitor and target–competitor conditions, respectively, were larger in the low-cuing group (11% inhibition and 10% facilitation) than they were in the high-cuing group (8% inhibition and 0% facilitation). It seems that if covert cuing had any effect, it was to mute both inhibition and facilitation. Thus, the inhibition and facilitation effects in the competitor–competitor and target–competitor conditions seem likely to reflect changes to features of the affected items.

In addition to asking about covert cuing strategies on the final test, our questionnaire asked whether subjects did extra retrieval practice during the practice phase. We included this question because we thought that subjects might adopt special strategies during the practice phase that might reduce retrieval-induced forgetting in the target–competitor condition artificially. For instance, after recalling a practiced item on a given retrieval-practice trial, subjects might have spent any extra time deliberately rehearsing the unpracticed competitor with which that practice target was

---

[4] Using an incidental encoding method also reliably reduced subjects' tendency to deliberately integrate items across exemplar pairs in Experiment 1 ($M = 2.79$) compared with the pilot study ($M = 3.38$), $F(1, 96) = 12.38$, $p < .001$, $MSE = 0.78$.

paired. Indeed, subjects did report doing such extra practice (M = 3.1 on a 5-point scale). However, there were no reliable differences in the degree to which unpracticed competitors were facilitated in the target–competitor condition across the low (mean rating = 1.70, facilitation = 7%), moderate (mean rating = 3.09, facilitation = 11%), and high extra-practice groups (mean rating = 4.44, facilitation = 3%). Thus, the facilitation of unpracticed competitors in the target–competitor condition is unlikely to have been caused by the deliberate rehearsal of those items during the retrieval-practice phase. Rather, facilitation is likely to have been caused by some more indirect process, such as the strengthening of shared features.[5]

Similarity ratings. Subjects reported generating averages of 0.99 and 1.27 similarities in the target–competitor and competitor–competitor conditions, respectively, as measured during the similarity rating phase. As an exploratory analysis, we correlated these ratings with the amount of inhibition that each subject exhibited. Although no strong relationships were found, the correlations were in the theoretically expected direction for the target–competitor (r = −.12) and competitor–competitor conditions (r = .07). These correlations may be weak because our rating scale did not allow subjects to report the full range of similarities they generated (0 = *none,* 1 = *one,* 2 = *more than 1*).

Experiment 2

In Experiment 1, unpracticed competitors were facilitated by retrieval practice when they were initially encoded together with to-be-practiced targets. If this reversal in retrieval-induced forgetting was caused by an increase in target–competitor similarity, as we have proposed, then if those same target–competitor pairs are encoded so as to make them less similar to one another, the retrieval-based facilitation found in Experiment 1 should revert to retrieval-induced forgetting. This reversal should arise for three reasons. First, the absence of instructions to find similarities should reduce the number of features shared by target and competitors compared with that present in Experiment 1. This should decrease the facilitatory influence of shared features. Second, any effort to find differences between targets and competitors should increase the number of nonoverlapping features that are susceptible to suppression. Finally, because targets and competitors are initially encoded as members of the same category, there should be sufficient overlap between their representations to ensure that they compete, despite whatever encoding operations may be directed at making them less similar.[6] In Experiment 2, we tested this prediction by replacing the similarity encoding instructions of Experiment 1 with instructions to find unique properties of each exemplar in a pair.[7] If target–competitor similarity decreases retrieval-induced forgetting, as suggested by the Anderson and Spellman (1995) model, distinctive encoding should resurrect retrieval-induced forgetting in Experiment 2.

Replacing similarity encoding with unique property encoding allows us to test predictions about competitor–competitor similarity as well. If competitor–competitor similarity increases retrieval-induced forgetting, as we have suggested, then asking subjects to find unique properties of items in competitor–competitor pairings should reduce retrieval-induced forgetting compared with that found in Experiment 1. This reduction should occur because the absence of similarity encoding instructions in Experiment 2 will diminish the number of features that competitors share with each other. This decrease in the number of overlapping features should reduce the chance that inhibiting one competitor will also affect the others. If retrieval-induced forgetting is reduced in the competitor–competitor condition of Experiment 2, it will suggest that the advantage of distinctive encoding observed by R. E. Smith and Hunt (in press) may derive more from competitor–competitor distinctiveness than from target–competitor distinctiveness. This conclusion would be even more likely if we also found that increasing target–competitor distinctiveness increased inhibition.

Method

Subjects

Thirty-two University of Oregon undergraduates participated to fulfill a course requirement.

---

[5] Unpracticed competitors in the target–competitor condition may also have been automatically primed by the successful retrieval of their paired retrieval-practice target. Such priming may have offset suppression of the unpracticed competitor's distinctive features, causing facilitation. We cannot rule out the contribution of such a process to the present data. However, compensatory priming is not incompatible with our present theoretical proposal.

[6] According to the pattern-suppression model, some amount of overlap between a target and a competitor is necessary for competition to ensue and thus for inhibition to be necessary. However, it is unclear exactly what should count as "nonoverlapping" items. If two members of a category (or two associates of a cue generally) had no obvious shared features, should they be considered unrelated (nonoverlapping)? Or does the fact that the items (regardless of how dissimilar they may otherwise be) share a retrieval cue constitute some minimum amount of similarity (overlap) that would form the basis of competition? Because there is a wealth of data from the classical interference literature showing that completely dissimilar responses to a shared retrieval cue nonetheless compete, and because encoding two items with respect to a common cue seems likely to generate at least some common properties, we consider nonoverlapping items to be those that share no obvious features in common and that are not associated to a common cue.

[7] We asked people to find "unique properties," rather than asking them to find differences more generally. This emphasis on unique properties was premised on the assumption that there may be a difference between "alignable" and "nonalignable" differences (Markman & Gentner, 1993). For instance, *Apples* and *Bananas* differ on the dimension of color, but they share the color dimension generally. This is an example of an alignable difference because the two items can be aligned along the shared dimension of color. Although generating "is yellow" in this case would be generating a difference between the items at the feature value level, it may actually increase similarity at the level of dimensions (by fostering the encoding of the color dimension), making it unclear whether the manipulation is actually decreasing or increasing similarity. Not all differences require the presence of a shared dimension, however. For instance, *Surfboard* and *Sailboat* are different in that sailboats have steering mechanisms and surfboards do not. This is a nonalignable difference because there is (arguably) no shared dimension along which the items differ in value. Thus, if subjects generate "has a steering mechanism" as a unique property of sailboats, they are increasing the differences between items without increasing their dimensional similarity. Thus, we asked subjects to find unique properties (using the foregoing example) of items in the hopes of creating a better manipulation of difference encoding. Whether alignable and nonalignable differences have a different effect on retrieval-induced forgetting remains to be established.

Design

The design of Experiment 2 was identical to Experiment 1, except that pattern-of-comparison was manipulated by asking subjects to find unique properties of exemplars rather than similarities.

Materials and Procedure

The materials and procedure of Experiment 2 were identical to those of Experiment 1, except for the new instructions given during the second encoding phase. Instead of asking subjects to find similarities between each pair of exemplars, we asked them to find unique properties. Specifically, for each pair, subjects were to identify as many distinctive characteristics of each exemplar as they could. As an example, we gave the triplet *Vehicle Surfboard Sailboat* and told subjects that the characteristic "has a steering wheel" is unique to sailboats and that the characteristic "can carry" is unique to surfboards. Subjects indicated the number of unique properties they generated by marking the same 3-point scale used to rate similarities in Experiment 1, with numbers ranging from 0 (*no unique properties found*) to 2 (*2 or more unique properties found*).

Results and Discussion

All of the analyses were performed with study and practice counterbalancing as between-subjects variables. An alpha level of .05, two-tailed, was adopted for all of the analyses.

Retrieval Practice

Retrieval-practice success rates did not vary across the competitor–competitor (M = 73%) and target–competitor (M = 74%) conditions ($F < 1$).

Final Recall Performance

As in Experiment 1, overall recall performance was better in the competitor–competitor condition (M = 49%) than in the target–competitor condition (M = 40%), $F(1, 24) = 12.33$, MSE = 203.76, again suggesting that comparing items within the cross-categorizable exemplar set enhances accessibility from the extralist category cue.

Effects of retrieval practice. The inhibition effect, averaged over pattern of comparison (unpracticed competitor − baseline = 41% − 48% = −7%) was highly significant, $F(1, 24) = 9.74$, MSE = 153.2. More important, however, was that the pattern of inhibition observed for the target–competitor condition in Experiment 2 was the opposite of that seen in Experiment 1 (see Figure 5b): Whereas similarity-based encoding facilitated unpracticed competitors in Experiment 1, encoding unique properties led to significant inhibition, $F(1, 24) = 10.56$, MSE = 154.35, as predicted by the Anderson and Spellman (1995) model. In contrast, encoding unique properties in the competitor–competitor condition was associated with very little retrieval-induced forgetting (unpracticed competitor − baseline = −4%)—an effect that was not reliable, $F(1, 24) = 1.48$, $p = .24$, MSE = 137.5, although the difference in the amount of inhibition across the target–competitor (10%) and competitor–competitor conditions (4%) was marginal, $F(1, 24) = 1.50$, $p = .13$, MSE = 138.7. Nevertheless, for the competitor–competitor condition, reliable inhibition was found with similarity-based encoding (Experiment 1) but not with the encoding of unique properties (Experiment 2).

Comparison with Experiment 1. To confirm the contrasting findings of Experiments 1 and 2, we incorporated the data from these studies into a single analysis. If target–competitor and competitor–competitor similarity have different effects on the amount of retrieval-induced forgetting, the way that the level of similarity (Experiment 1 vs. Experiment 2) affects the amount of inhibition (unpracticed competitor vs. baseline) should differ reliably as a function of pattern-of-comparison (target–competitor vs. competitor–competitor). Consistent with our hypothesis, this three-way interaction was highly significant, $F(1, 64) = 12.82$, MSE = 0.021. This difference in the effects of similarity on inhibition between the two patterns of comparison was brought about by two changes, consistent with our hypotheses. First, in the target–competitor condition, there was a significant reduction in retrieval-induced forgetting (a reversal, in fact) when similarity was high (Experiment 1) compared with when it was low (Experiment 2), $F(1, 64) = 17.89$, MSE = 0.016. Second, in the competitor–competitor condition, there was a weak trend toward more inhibition (baseline − unpracticed competitor) when similarity was high (Experiment 1) than when it was low (Experiment 2), $F(1, 64) = 1.81$, $p = .18$, MSE = 0.022, again consistent with predictions. Although this latter Inhibition × Experiment interaction is not reliable, the pattern in these competitor–competitor data is identical to that found by R. E. Smith and Hunt (in press)—significant inhibition with similarity-based encoding and no significant inhibition for distinctiveness encoding. Because R. E. Smith and Hunt reported separate tests of inhibition in the similarity and distinctiveness encoding conditions of their study without ever testing their interaction, the present finding can be considered, by comparison with their reported data, as a replication of their findings. Nevertheless, the overall effects of similarity on retrieval-induced forgetting differ reliably across target–competitor and competitor–competitor conditions.

Postexperimental questionnaire analysis. The pattern of differential impairment evident in the overall data appears more robust for subjects who reported that they did not engage in special covert cuing strategies on the final test (1% facilitation and 14% inhibition in the competitor–competitor and target–competitor conditions, respectively), which is supported by a significant interaction of inhibition with pattern of comparison for this low-cuing group, $F(1, 64) = 4.10$, MSE = 217.3. Subjects reporting a high degree of covert cuing, as in Experiment 1, showed a muted version of this effect (8% competitor–competitor inhibition [compared with 10% in Experiment 1] and 7% target–competitor inhibition [compared with 7% facilitation in Experiment 1]), although the three-way interaction of inhibition, pattern of comparison, and degree of covert cuing did not reach statistical significance, $F(1, 64) = 2.60$, $p = .11$, MSE = 217.3. It is unclear why high covert cuing might diminish the effects in this way. Nevertheless, use of a covert cuing strategy clearly cannot explain the overall pattern of inhibition found in this experiment—if anything, this strategy weakens the overall pattern, as in Experiment 1.

The degree of extra retrieval practice reported on the questionnaire (low vs. high) did not modulate the amount of competitor–competitor (2% vs. 5% in low vs. high) or target–competitor inhibition (10% vs. 11%).

Difference ratings. The mean number of differences generated was 1.83 and 1.65 for the target–competitor and competitor–competitor conditions, respectively. As was the case for the similarity rating measure in Experiment 1, no strong correlations were observed between these ratings and inhibition for the target–competitor condition (r = .18) or for the competitor–competitor condition (r = .04).

General Discussion

Previous studies have shown that retrieving an item from long-term memory can impair the retention of related competitors. Experiments 1 and 2 of the present study demonstrate that the degree to which a competing memory will be impaired by a target item's retrieval depends on how related it is to the target, and also to other competitors in memory. When a competitor is very similar to a target, it will be less impaired by the target's retrieval than when it is less similar but still related. On the other hand, when competing memories are very similar to one another, they will be more impaired by the target's retrieval than when they are dissimilar to each other. Thus, target–competitor and competitor–competitor similarity have opposite effects on the amount of retrieval-induced forgetting. These effects of similarity on retrieval-induced forgetting are unlikely to arise from the use of associative retrieval paths linking targets to competitors; rather, they are more likely to arise from changes to the items themselves. Three findings support these conclusions.

First, when subjects encoded category members by finding similarities between to-be-practiced targets and competitors (the target–competitor condition), later retrieval practice on target items actually facilitated delayed recall of the competitors rather than causing retrieval-induced forgetting. This facilitation reverted to significant retrieval-induced forgetting in Experiment 2 when subjects were asked to find distinctive properties of each item in a target–competitor pair. This reversal was observed even though subjects in Experiment 2 were presented with exactly these same items for the same amount of time and even though they performed an encoding task that required concurrent processing of both items. Thus, the main factor determining whether unpracticed competitors were facilitated or impaired by later retrieval practice on target items was the degree of target–competitor similarity fostered by the encoding task.

Second, when subjects looked for similarities between the competitors themselves during encoding, those competitors were suppressed by later retrieval practice on the target items. This retrieval-induced forgetting was reduced in Experiment 2, in which subjects looked for distinctive properties of the competitors. Here again, this reduction in impairment was found even though subjects in Experiment 2 were presented with exactly the same items for the same amount of time and had to process the items together. Thus, more retrieval-induced forgetting will be observed when competitors are similar to one another than when they are dissimilar. Together with the findings of the target–competitor condition, these data argue that target–competitor and competitor–competitor similarity have opposite effects on the amount of retrieval-induced forgetting that will be observed.

Third, and finally, the present differences in retrieval-induced forgetting occurred even though the unpracticed competitors were tested with novel extralist category cues. When subjects tried to recall the unpracticed competitor *Red Tomato* with the cues *Food T—*, they did so more poorly when they had practiced *Red Heart* earlier on than when they did not. That impairment of *Tomato* generalized to retrieval cues other than those used to perform retrieval practice replicates earlier work showing that retrieval-induced forgetting is cue independent (Anderson & Spellman, 1995). Cue independence is usually taken as evidence that the representation of the unpracticed competitor itself was impaired by retrieval practice rather than any particular associative route into that item. In the present study, using an independent retrieval cue during the test phase also allowed us to establish that variations in the amount of retrieval-induced forgetting caused by our similarity manipulations reflected feature-level changes to the affected items themselves. Because our independent test cues (e.g., *Food*) were designed to be unrelated to the practiced targets (e.g., *Red Heart*), any associations that might exist between targets and unpracticed competitors (e.g., between *Heart* and *Tomato*) were unlikely to have been available when subjects tried to recall that competitor (e.g., *Food T*—). Thus, the present variations in impairment caused by target–competitor similarity are unlikely to reflect the benefits of having associative connections between unpracticed competitors and retrieval-practice targets.

Although our findings support the prediction that competitor–competitor distinctiveness reduces retrieval-induced forgetting, an alternative account of these data cannot be ruled out. In the competitor–competitor conditions of Experiments 1 and 2, subjects were asked to think of similarities or differences between pairs of to-be-practiced items as well as pairs of competitors. Both the to-be-practiced targets and the competitors were encoded in this way to match the conditions present for categories in the target–competitor condition, in which all of the exemplars of a category took part in the similarity (or difference) encoding phase. Because subjects compared pairs of practiced items as well as pairs of competitors, however, one might argue that target–target distinctiveness reduced retrieval-induced forgetting in Experiment 2 instead of competitor–competitor distinctiveness. This suggestion is reasonable if one considers the fact that practice targets suffer competition not only from unpracticed competitors but also from other practice targets. Making targets more distinct from each other might thus reduce competition exerted by other retrieval-practice targets during the practice of a given item. If one assumes that reducing competition in this fashion lowers the overall need for inhibition, perhaps unpracticed competitors might be less suppressed as a result, causing less retrieval-induced forgetting. Thus, increasing target–target distinctiveness might have reduced retrieval-induced forgetting in the competitor–competitor condition of Experiment 2.

One problem with this view is that it makes assumptions about the effects of distinctiveness that contradict findings from our target–competitor conditions. Key to this alternative is the idea that a retrieval-practice target functionally serves as a competitor whenever another target is being practiced. If the targets themselves serve as competitors, then manipulations that increase target–target distinctiveness are, in effect, increasing target–competitor distinctiveness. For an increase in target–target distinctiveness to reduce retrieval-induced forgetting, then, we must assume that target–competitor distinctiveness reduces retrieval-induced forgetting. This assumption contradicts the findings of the target–competitor condition in Experiment 2: Target–competitor distinctiveness served to increase retrieval-induced forgetting, not reduce it. Thus, target–target distinctiveness seems unlikely to have reduced retrieval-induced forgetting in the competitor–

competitor condition of Experiment 2. Nevertheless, the contribution of this factor to the present competitor–competitor similarity effect should be assessed empirically in future work.

The differing effects of target–competitor and competitor–competitor similarity on retrieval-induced forgetting testify to the importance of the pattern of interitem similarity in determining the amount of retrieval-induced forgetting that will be observed. When there are many competitors in memory, one must consider not only whether they are similar to the retrieval target but also to each other. If this hypothesis is correct, it may help to explain why previous attempts to understand the relation between similarity and retrieval-induced forgetting have yielded inconsistent findings. In the next section, we consider the relevance of our findings to interpreting the data of R. E. Smith and Hunt (in press) and Hartinger and Bauml (1999). We then discuss the implications of this proposal for work on the similarity paradox in classical studies of interference (Osgood, 1949). We close by discussing two theoretical approaches to the present findings: the center–surround approach to semantic retrieval (Carr & Dagenbach, 1990) and the pattern-suppression model (Anderson & Spellman, 1995).

Similarity and Retrieval-Induced Forgetting

Previous studies of retrieval-induced forgetting have found that interitem similarity both increases (R. E. Smith & Hunt, in press) and decreases (Hartinger & Bauml, 1999) impairment of unpracticed competitors. One approach to interpreting these findings is to assume that each conclusion is correct but for different reasons. Examination of the procedures for manipulating similarity in these two studies reveals that they may have placed differing emphases on target–competitor and competitor–competitor similarity. In Hartinger and Bauml's study, the similarity manipulation focused mainly on target–competitor similarity. In their procedure, subjects studied four items per category and did retrieval practice on only one item that was either from the same subcategory as a critical unpracticed competitor or from a different subcategory. When the practiced item shared a subcategory with the unpracticed competitor, retrieval-induced forgetting was reduced. In our Experiment 1, subjects encoded target–competitor similarities and retrieval-induced forgetting was reduced. These findings confirm Hartinger and Bauml's conclusions about similarity and retrieval-induced forgetting and extend their findings to a procedure in which target–competitor similarity is manipulated through encoding instructions alone. Our findings also suggest that Hartinger and Bauml's reduction in retrieval-induced forgetting was not caused solely by retrieval strategies using associative mediation from practiced items to unpracticed competitors on the final test, because we found the same reduced impairment with an independent retrieval cue.

In R. E. Smith and Hunt's (in press) study, however, interitem similarity increased retrieval-induced forgetting. More impairment was found when their subjects were asked to make every exemplar in a category similar to every other than when they were asked to make every exemplar different. Unlike in Hartinger and Bauml's (1999) study, R. E. Smith and Hunt's procedure encouraged the encoding of both target–competitor and competitor–competitor similarities. These additional variations in competitor–competitor similarity may explain the discrepancy between their findings and those of Hartinger and Bauml. When our subjects encoded competitor–competitor similarities, they showed more retrieval-induced forgetting than when they encoded competitor–competitor differences. It seems possible that the reduced impairment in R. E. Smith and Hunt's distinctive encoding condition may have been caused more by competitor–competitor distinctiveness than by target–competitor distinctiveness. This possibility is reinforced by the fact that increasing target–competitor distinctiveness in our Experiment 2 increased retrieval-induced forgetting. Although this account of R. E. Smith and Hunt's data is different from theirs, it agrees with their proposition that distinctive encoding may, in some cases, protect against retrieval-induced forgetting.

Although the foregoing analysis seems promising as a reconciliation of these findings, ambiguities remain. First, even though R. E. Smith and Hunt's (in press) procedure encouraged competitor–competitor similarities, the target–competitor and competitor–competitor similarities that subjects found were not independent. The encoding instructions asked subjects to find a feature that made a given item similar to every other. Thus, the features that made competitors similar to one another also would have been the features that overlapped with practiced targets. These features should have been strengthened by retrieval practice and not suppressed and thus could not have been responsible for the increased inhibition in the similarity encoding condition. It seems likely, however, that in trying to think of features that an item might share with all of its category mates, subjects would have tried and rejected many attributes that matched only a few other items. For example, before discovering that *Cherry, Raspberry,* and *Apple* are all red, subjects might have first tried the feature "berry" and rejected it because it was not shared by *Apple.* The increased encoding of these partial similarities seems likely to have generated independent competitor–competitor similarities that could have made the unpracticed competitors more vulnerable to retrieval-induced forgetting. Second, even if R. E. Smith and Hunt's similarity encoding procedure produced independent competitor–competitor similarities, it is unclear how the influence of that factor should be expected to combine with the effects of target–competitor similarity also generated by their procedure.

The issue of how the effects of target–competitor and competitor–competitor similarity combine to determine the amount of retrieval-induced forgetting is general, and a resolution of it is likely to be important to interpreting the results of other studies as well. One possibility is that these factors combine additively, with the amount of impairment determined by the proportions of these types of similarity encoded by subjects. If the proportions of target–competitor and competitor–competitor similarity vary randomly from study to study, it will be difficult to predict whether manipulations of overall item similarity should increase, decrease, or leave unaffected the amount of impairment. The contributions of these opposing factors may explain why attempts to find a relation between intercategory similarity and output interference have found no reliable differences in impairment (Roediger & Schmidt, 1980).[8] If output interference is produced by the same

[8] Although Roediger and Schmidt (1980) found no reliable differences between the amount of output interference observed for similar and dissimilar categories, it should be noted that (a) there was a numerical difference in favor of similar categories showing more output interference (comparing the recall levels of the first and last categories in their output sequence) and (b) the similar category condition showed much lower

mechanisms as long-term retrieval-induced forgetting, which seems likely (Anderson et al., 1994; Anderson & Neely, 1996; Bauml, 1998), then having subjects study highly similar categories should produce both the positive effects of target–competitor similarity as well as the negative effects of competitor–competitor similarity. The net result of these factors may have been to obscure systematic relations between these two types of similarity and output interference. Whatever the proper explanation of these findings may be, the present study recommends that the influence of these factors be taken into account in any analysis of the relation between similarity and retrieval-induced forgetting.

Although the present theory can explain some persisting inconsistencies in the literature on similarity and inhibition (see the next section), one might be concerned that the theory is too flexible. Given two dimensions of similarity that have opposing effects on the amount of inhibition that will be observed, it might seem that one could explain any function relating item similarity to inhibition. For instance, to explain increases or decreases in impairment with interitem similarity, one need only assume different proportions of target–competitor and competitor–competitor similarity for the materials in a given experiment. Although concern over this flexibility is warranted, we argue that the difficulty lies not so much in the theory as it does in the paradigms often used to study the relationship between similarity and inhibition. Many paradigms such as free or cued recall of categorized word lists (e.g., Roediger & Schmidt, 1980; Shuell, 1968) cannot provide clear tests of the theory because (a) when subjects are free to recall items in any order, one cannot specify in advance what counts as target–competitor and competitor–competitor similarity and, consequently, (b) the degree of target–competitor and competitor–competitor similarity cannot be separately measured or manipulated. A test of the theory requires paradigms that separate these dimensions and that provide a way of varying similarity independently within each dimension. Given these constraints, one should be able to isolate target–competitor and competitor–competitor similarity, as we have done in the present study. The present dissociation suggests that there are both theoretical and empirical reasons to prefer a more complex account of how similarity affects inhibition.

The Similarity Paradox in Classical Interference Research

Many classical studies of retroactive interference used procedures that encouraged retrieval-based suppression (see Anderson et al., 1994; Anderson & Neely, 1996; and Bauml, 1996, for discussions). It is thus not surprising that some of the same complexities that have arisen in current work on retrieval-induced forgetting also characterized early work on retroactive interference. For instance, many early interference studies found that at very high levels, interlist similarity decreased retroactive interference (Cheng, 1929; Dreis, 1933; Harden, 1929; Kennelly, 1941; Robinson, 1927; Watson, 1938). Yet, in other studies, there was compelling evidence that the amount of retroactive interference suffered by a first study list actually increased with the degree of similarity between it and a second study list (McGeoch & McDonald, 1931; McGeoch & McGeoch, 1936). In a classic article, Osgood (1949) noted that the empirical generalization promoted by McGeoch and McDonald's findings led to a paradox: If retroactive interference truly increases with the degree of interlist similarity, then the greatest amount of retroactive interference should be found when a second list is identical (i.e., maximally similar) to the first study list. This implication is at odds with the compelling fact that people's memories get better with repetition, not worse. Osgood noted that this paradox, along with the contradictions between these classical findings, could be reconciled if a distinction was drawn between the degree of stimulus and response similarity present between the first and second study lists.

According to Osgood's (1949) proposal, the amount of retroactive interference suffered by a first list of paired associates (e.g., pairs like *Dog Rock*) should increase as the *stimulus members* of the second list become similar to those on the first list (i.e., as they approach identity to the first list; e.g., *Dog Cloud*) but should decrease as the *response members* on the second list become similar to those on the first list (e.g., *Dog Stone*). This emphasis on stimulus similarity in producing retroactive interference fits with studies that found greater retroactive interference when the stimulus terms from the two study lists were identical than when they were different. By this view, previous studies showing that interference increased with interlist similarity could be attributed to stimulus similarity across lists. Compatible with this idea, most earlier studies showing that interlist similarity increased retroactive interference used serial learning procedures in which stimulus and response similarity could not be easily disentangled (McGeoch & McDonald, 1931; McGeoch & McGeoch, 1936). However, when stimulus similarity was held constant (at identity) and response similarity was varied across study lists, Osgood (1946, 1948) showed that retroactive interference actually decreased with interresponse similarity, consistent with his proposal and with the notion of response generalization (e.g., Underwood & Hughes, 1950). Many studies since then have replicated the reduction in retroactive interference with response similarity (e.g., Dallett, 1962; Kanungo, 1967; Morgan & Underwood, 1950; Postman, 1964; Young, 1955), even when the second list of pairs was given as many as 20 learning trials (Barnes & Underwood, 1959).

Like Osgood's (1949) resolution to the classical similarity paradox, we offer a distinction between two kinds of similarity to account for otherwise contradictory findings on retrieval-induced forgetting. Rather than being an alternative framework, however, the present proposal should be viewed as an extension of Osgood's classic idea. Osgood's framework was built around the retroactive interference design, in which there was typically only one target item (the List 2 response to a stimulus) and one competitor (the List 1 response to that same stimulus). Because of this design, Osgood's analysis never needed to account for anything beyond target–competitor similarity. His analysis remains useful today because it captures the wealth of classical data showing that response similarity reduces retroactive interference and also because it is compatible with the present target–competitor findings. Our competitor–competitor similarity results suggest that his response similarity concept was too simple, however. If Osgood's analysis is to extend to cases in which there are many competitors

overall recall than did the dissimilar category condition. It is possible that if the recall data were expressed in terms of proportion loss from the recall level of the first category in the testing sequence (to correct for baseline differences), output interference might have been reliably greater in the similar category condition.

in memory, response similarity needs to be broken into both target–competitor and competitor–competitor similarity because these factors have different effects on the amount of inhibition that will be observed.

Theoretical Approaches

Osgood's (1949) classical framework captures important generalizations regarding the effects of similarity on interference and inhibition. His framework is mainly descriptive, however. In this section, we describe two theoretical approaches to retrieval that provide mechanisms that can explain aspects of the present data: the center–surround theory (Carr & Dagenbach, 1990) and the pattern-suppression model (Anderson & Spellman, 1995).

Center–Surround Theory

Although the present study was motivated by the pattern-suppression model, the center–surround theory of Carr and Dagenbach (1990) can also explain aspects of our findings. The center–surround theory posits that when retrieval of a word's meaning is extremely difficult (either because the meaning is weakly learned or because the perceptual input supposed to cue that meaning is degraded), attention is so centered on the representation of the sought-after meaning that the representations of other words related to it are actively suppressed. The function of attention in this situation is twofold—to activate that area of semantic space that is the retrieval focus (i.e., the semantic "center," or the target item's meaning) and to inhibit nearby regions in semantic space that impede the target's retrieval (i.e., the "surround," or the competitors' meanings). Thus, when weak codes need to be discriminated, attentional processes provide top-down input that helps separate the target from distractors, taking advantage of lateral inhibitory mechanisms thought to be analogous to center–surround networks in perceptual systems.

Like the pattern-suppression model, the center–surround theory posits that retrieving a target from long-term memory can suppress similar, competing representations. Although this view might seem to predict more retrieval-induced forgetting as items become highly similar, the prediction depends on the degree of similarity between the target and the competitor. If a competitor substantially overlaps the target in semantic space, enough of its representation may be in the attentional center so as to create facilitation rather than inhibition. This implication was tested recently by Barnhardt, Glisky, Polster, and Elam (1996), who showed that a very difficult retrieval of the meaning of an uncommon word on a prime trial actually facilitated a subsequent lexical decision for a direct synonym of the prime, even though it slowed lexical decision times to semantically similar nonsynonyms. Barnhardt et al. argued that the facilitation of synonyms provides strong support for the center–surround theory because synonyms should occupy a region in semantic space that overlaps with the target, preventing them from being suppressed.

Center–surround theory can be adapted to provide a partial explanation for the effects of similarity on retrieval-induced forgetting in the present studies. In particular, the reduction in retrieval-induced forgetting associated with target–competitor similarity in Experiment 1 may be produced by mechanisms similar to those thought to be at work in the study of Barnhardt et al. (1996). If we assume that encoding target–competitor similarities brings a target and a competitor closer together in semantic space, some portion of the unpracticed competitor's representation might fall within the attentional center of the item being recalled during retrieval practice. If the item in the attentional center gets strengthened by its retrieval, then retrieval practice might actually facilitate unpracticed competitors if those competitors overlap with the practice target. Thus, the center–surround theory could explain our target–competitor similarity results. It is less clear how the model would account for the effects of competitor–competitor similarity. Clear predictions about this factor are made by the pattern-suppression model, which we discuss next.

Pattern-Suppression Model

Like the center–surround theory, the pattern-suppression model (Anderson & Spellman, 1995) posits a focus of activation on the target item and a need to suppress highly similar competing representations in order to discriminate the target. However, the pattern-suppression model is more specific in its commitment to distributed representations. This feature of the model enables it to predict the effects of both target–competitor and competitor–competitor similarity on retrieval-induced forgetting with a simple principle: Any effect on a unit (i.e., inhibition, facilitation) for one memory should also affect the other memories in which that unit takes part. This principle implies that target–competitor and competitor–competitor similarity should have opposite effects on retrieval-induced forgetting because shared feature units are facilitated when a competitor is similar to a target but are inhibited when a unit is shared by two competitors (and not the target). These findings lend strong support to the pattern-suppression approach and demonstrate the importance of considering the fate of shared features in predicting how similarity affects retrieval-induced forgetting.

Although the pattern-suppression model predicts that retrieval-induced forgetting should be reduced at very high levels of target–competitor similarity, the theory does not always predict that target–competitor similarity should reduce impairment. In fact, the relation between target–competitor similarity and inhibition should be nonmonotonic. For instance, we know that complete dissimilarity between a practice target and a competitor (e.g., encoding under different categories) yields little retrieval-induced forgetting. As target–competitor similarity is brought from this extremely low level to an intermediate level, however, retrieval-induced forgetting should increase. This follows because the model assumes, as do R. E. Smith and Hunt (in press) and Carr and Dagenbach (1990), that suppression is necessary when a target and a competitor become hard to discriminate. Some amount of similarity must be present for inhibition to be needed. Where the model differs from R. E. Smith and Hunt's analysis is in its focus on shared feature facilitation as a compensatory factor that, at high levels of similarity, ought to make the amount of behavioral impairment run counter to the actual amount of suppression that must be exerted. Unfortunately, this function makes predictions about similarity and inhibition difficult. Without knowing where a pair of conditions falls on the target–competitor similarity continuum, one cannot know whether to expect inhibition to increase with similarity (e.g., when the points fall on the lower end), to not change (when the points are in the middle), or to decrease (e.g., when the points are on the high end). However, it should be possible to parametrically manipulate the level of similarity from complete dissimilarity to identity and observe the reversal of inhibition at some intermediate point.[9]

The way in which the pattern-suppression model predicts the effects of competitor–competitor similarity can be compared with its ability to explain the unusual finding of second-order inhibition, the phenomenon that inspired the model initially (Anderson & Spellman, 1995). In several experiments, Anderson and Spellman found that retrieval-induced forgetting observed in one category often generalized to members of other categories that were not at all similar to the items that received retrieval practice. For instance, in Experiment 2, they found that performing retrieval practice on *Green Emerald* not only impaired subjects' final recall of *Green Lettuce, Emerald*'s direct competitor, but also the recall of *Soups Mushroom*—an item that was similar to *Lettuce* but not itself similar to *Emerald.* Anderson and Spellman (1995) noted that this *second order inhibition* could be explained if *Mushroom* overlapped in semantic feature units with *Lettuce,* the direct recipient of suppression during the retrieval practice of *Green Emerald.* By sharing feature units with a competitor to the retrieval target, *Mushroom* should also be suppressed as well, even though *Mushroom*'s representation may not overlap with that of the target. The present account of the effects of competitor–competitor similarity makes use of the same principle of shared feature suppression but applies the principle to several competitors that are each directly related to the retrieval-practice target.

Concluding Remarks

Previous work has shown that retrieving a memory can impair the long-term retention of semantically related events. The present work shows how the negative effects of retrieval can be attenuated or magnified by similarity relationships between the retrieval target and competing memories or between the competing memories themselves. When a competitor is very similar to a retrieval target, the negative effects of inhibition may be less apparent than when competitors are less similar. Such reductions in impairment should occur to the extent that facilitation of the features of the correctly recalled target memory generalizes to those competitors that share those feature units. Facilitation of shared features may compensate for suppression of the distinctive ones, preserving access to memories that might otherwise be rendered inaccessible. Thus, when it comes to target–competitor similarity, whatever is not inhibited about a competitor can make it stronger.

When competitors are similar to each other, however, the behavioral effects of suppression will be magnified to the extent that the suppression of the features of one competitor generalizes to others that share those feature units. Thus, when many similar experiences compete, the impact of suppression on those competing memories may be particularly pronounced. Although target–competitor and competitor–competitor similarity appear to behave very differently, we have suggested that these differing effects reflect a common underlying principle: Changes that affect the features of one pattern (i.e., activation, inhibition) will affect in similar ways all other patterns in which those feature units participate. Whatever the correct account of these similarity effects may be, the present findings illustrate the important role of representation in determining the behavioral impact of the inhibitory processes that support memory retrieval.

[9] A similar nonmonotonic function relating similarity and interference was proposed in the early days of the classical interference era and was known as the Skaggs–Robinson hypothesis (Robinson, 1927; Skaggs, 1925, 1927). The Skaggs–Robinson hypothesis also noted that completely dissimilar and identical stimuli should not suffer impairment, whereas items with intermediate similarity should. However, the Skaggs–Robinson hypothesis did not distinguish target–competitor from competitor–competitor similarity. Thus, it is unclear whether attempts to test this hypothesis have direct bearing on the present prediction, given that care would not have been taken to separate out the opposing influences of these factors.

References

Allen, G. A., Mahler, W. A., & Estes, W. K. (1969). Effects of recall tests on long-term retention of paired associates. *Journal of Verbal Learning and Verbal Behavior, 8,* 463–470.

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (in press). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic Bulletin and Review.*

Anderson, M. C., & Bjork, R. A. (1994). Mechanisms of inhibition in long-term memory: A new taxonomy. In D. Dagenbach & T. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 265–325). San Diego, CA: Academic Press.

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,* 1063–1087.

Anderson, M. C., & McCulloch, K. C. (1999). Integration as a general boundary condition on retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 608–629.

Anderson, M. C., & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Memory: Handbook of perception and cognition* (2nd ed., pp. 237–313). San Diego, CA: Academic Press.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review, 102,* 68–100.

Barnes, J. M., & Underwood, B. J. (1959). "Fate" of first-list associations in transfer theory. *Journal of Experimental Psychology, 58,* 97–105.

Barnhardt, T. M., Glisky, E. L., Polster, M. R., & Elam, L. (1996). Inhibition of associates and activation of synonyms in the rare-word paradigm: Further evidence for a center-surround mechanism. *Memory & Cognition, 24,* 60–69.

Bauml, K. (1996). Revisiting an old issue: Retroactive interference as a function of the degree of original and interpolated learning. *Psychonomic Bulletin and Review, 3,* 380–384.

Bauml, K. (1997). The list strength effect: Strength-dependent competition or suppression? *Psychonomic Bulletin and Review, 4,* 260–264.

Bauml, K. (1998). Strong items get suppressed, weak items do not: The role of item strength in output interference. *Psychonomic Bulletin and Review, 5,* 459–463.

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: Erlbaum.

Blaxton, T. A., & Neely, J. H. (1983). Inhibition from semantically related primes: Evidence of a category-specific retrieval inhibition. *Memory & Cognition, 11,* 500–510.

Brown, A. S. (1981). Inhibition in cued retrieval. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 204–215.

Carr, T. H., & Dagenbach, D. (1990). Semantic priming and repetition priming from masked words: Evidence for a center-surround attentional mechanism in perceptual recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 341–350.

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory & Cognition, 20,* 633–642.

Cheng, N. Y. (1929). Retroactive effect and degree of similarity. *Journal of Experimental Psychology, 12,* 444–449.

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

Crowder, R. G. (1976). *Principles of learning and memory.* Hillsdale, NJ: Erlbaum.

Dagenbach, D., Carr, T. H., & Barnhardt, T. M. (1990). Inhibitory semantic priming of lexical decisions due to failure to retrieve weakly activated codes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 328–340.

Dallett, K. M. (1962). The role of response similarity in proactive inhibition. *Journal of Experimental Psychology, 64,* 364–372.

Dreis, T. A. (1933). Two studies in retroaction: I. Influence of partial identity. II. Susceptibility to retroactive at various grade levels. *Journal of General Psychology, 8,* 157–171.

Gardiner, J. M., Craik, F. I. M., & Bleasdale, F. A. (1973). Retrieval difficulty and subsequent recall. *Memory & Cognition, 1,* 213–216.

Harden, L. M. (1929). A quantitative study of the similarity factor in retroactive inhibition. *Journal of General Psychology, 2,* 421–430.

Hartinger, A., & Bauml, K. (1999). Retrieval-induced forgetting: The role of item similarity. *Abstracts of the Psychononic Society, 4,* 57.

Hogan, R. M., & Kintsch, W. (1971). Differential effects of study and test trials on long-term recognition and recall. *Journal of Verbal Learning and Verbal Behavior, 10,* 562–567.

Kanungo, R. (1967). Meaning mediation in verbal transfer. *British Journal of Psychology, 58,* 205–212.

Kennelly, T. W. (1941). The role of similarity in retroactive inhibition. *Archives of Psychology, 37.*

Koutstaal, W., Schacter, D. L., Johnson, M. L., & Galluccio, L. (1999). Facilitation and impairment of event memory produced by photograph review. *Memory & Cognition, 27,* 478–493.

Kucera, H., & Francis, W. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Macrae, C. N., & MacLeod, M. D. (1999). On recollections lost: When practice makes imperfect. *Journal of Personality and Social Psychology, 77,* 463–473.

Markman, A. B., & Gentner, D. (1993). Splitting the differences: A structural alignment view of similarity. *Journal of Memory and Language, 32,* 517–535.

McGeoch, J. A., & McDonald, W. T. (1931). Meaningful relation and retroactive inhibition. *American Journal of Psychology, 43,* 579–588.

McGeoch, J. A., & McGeoch, G. O. (1936). Studies of retroactive inhibition: VI. The influence of the relative serial positions of interpolated synonyms. *Journal of Experimental Psychology, 19,* 1–23.

Morgan, R. L., & Underwood, B. J. (1950). Proactive inhibition as a function of response similarity. *Journal of Experimental Psychology, 40,* 592–603.

Osgood, C. E. (1946). Meaningful similarity and interference in learning. *Journal of Experimental Psychology, 36,* 277–301.

Osgood, C. E. (1948). An investigation into the causes of retroactive interference. *Journal of Experimental Psychology, 38,* 132–154.

Osgood, C. E. (1949). The similarity paradox in human learning: A resolution. *Psychological Review, 56,* 132–143.

Postman, L. (1964). Studies of learning to learn: Changes in transfer as a function of practice. *Journal of Verbal Learning and Verbal Behavior, 3,* 437–447.

Postman, L. (1971). Transfer, interference and forgetting. In J. W. Kling & L. A. Riggs (Eds.), *Woodworth and Schlosberg's: Experimental psychology* (3rd ed., pp. 1019–1132). New York: Holt, Rinehart & Winston.

Postman, L., & Underwood, B. J. (1973). Critical issues in interference theory. *Memory & Cognition, 1,* 19–40.

Robinson, E. S. (1927). The "similarity" factor in retroaction. *American Journal of Psychology, 39,* 297–312.

Roediger, H. L. III, & Neely, J. H. (1982). Retrieval blocks in episodic and semantic memory. *Canadian Journal of Psychology, 36,* 213–242.

Roediger, H. L. III, & Schmidt, S. R. (1980). Output interference in the recall of categorized and paired associate lists. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 91–105.

Shaw, J. S., Bjork, R. A., & Handal, A. (1995). Retrieval-induced forgetting in an eyewitness-memory paradigm. *Psychonomic Bulletin and Review, 2,* 249–253.

Shuell, T. J. (1968). Retroactive inhibition in free-recall learning of categorized lists. *Journal of Verbal Learning and Verbal Behavior, 7,* 797–805.

Skaggs, E. B. (1925). Further studies in retroactive inhibition. *Psychological Monographs, 34,* No. 161.

Skaggs, E. B. (1927). *The major forms of inhibition in man.* Chicago: University of Chicago Press.

Smith, A. D. (1971). Output interference and organized recall from long-term memory. *Journal of Verbal Learning and Verbal Behavior, 10,* 400–408.

Smith, R. E., & Hunt, R. R. (in press). The influence of distinctive processing on retrieval-induced forgetting. *Memory & Cognition.*

Solso, R. L., & Juel, C. L. (1980). Positional frequency and versatility of bigrams for two- through nine-letter English words. *Behavior Research Methods & Instrumentation, 12,* 297–343.

Underwood, B. J., & Hughes, R. H. (1950). Gradients of generalized verbal responses. *American Journal of Psychology, 63,* 422–430.

Watson, B. (1938). The similarity factor in transfer and inhibition. *Journal of Educational Psychology, 29,* 145–157.

Young, R. K. (1955). Retroactive and proactive effects under varying conditions of response similarity. *Journal of Experimental Psychology, 50,* 113–117.

Received September 13, 1999
Revision received March 31, 2000
Accepted March 31, 2000 ■