Journal of Experimental Psychology:
Learning, Memory, and Cognition
1999, Vol. 25, No. 3, 608–629

0278-7393/99/$3.00

Integration as a General Boundary Condition on Retrieval-Induced Forgetting

Michael C. Anderson and Kathleen C. McCulloch
University of Oregon

When people form connections between several memories that share a common retrieval cue, the tendency for those memories to interfere with one another during later retrieval attempts is often eliminated. Three experiments examined whether forming such connections might also protect memories from retrieval-induced forgetting, the phenomenon in which retrieving some associates of a cue leads to the suppression of others that interfere during retrieval (M. C. Anderson, E. L. Bjork, & R. A. Bjork, 1994). All 3 experiments found that instructing subjects to interrelate category exemplars during an initial study phase reduced retrieval-induced forgetting. Postexperimental questionnaires indicated that even when people were not instructed to interrelate exemplars, they often did so spontaneously and that this spontaneous integration also protected people from impairment. These findings, together with others obtained in different experimental settings, suggest that complex knowledge structures composed of highly interconnected components may be especially resistant to retrieval-induced forgetting.

Retrieval processes play a powerful role in shaping what we remember of our past. An act of recall may strengthen or impede memory for an event, depending on the circumstances. Remembering the event itself will enhance its retention, because retrieval is known to be a potent learning event (Allen, Mahler, & Estes, 1969; Bjork, 1975; Carrier & Pashler, 1992; Gardiner, Craik, & Bleasdale, 1973; Hogan & Kintsch, 1971). But retrieval plays another more negative role in shaping accessibility—a role intimately related to the classical phenomenon of interference (McGeoch, 1936; Melton & Irwin, 1940; for reviews, see Anderson & Neely, 1996; Crowder, 1976; Postman, 1971).

A growing body of research has shown the very act of remembering a past event can cause forgetting of related experiences. This *retrieval-induced forgetting* (Anderson, Bjork, & Bjork, 1994) can last at least 20 min and appears to be quite general. It has been induced by semantic retrieval (Blaxton & Neely, 1983; Dagenbach, Carr, & Barnhardt, 1990) as well as episodic recall (Anderson et al., 1994; Anderson & Spellman, 1995; Roediger, 1978; Shaw, Bjork, & Handal, 1995) and has been found with both verbal and visuospatial materials (Ciranni & Shimamura, 1998; see also Schooler, Fiore, & Brandimonte, 1997 for related effects). Retrieval-induced forgetting is believed to arise from the need to overcome interference from competing memories during retrieval. Interference is thought to trigger an inhibitory process that suppresses the competing item, ultimately causing it to be forgotten (Anderson et al., 1994; Anderson & Spellman, 1995; Anderson & Neely, 1996). The existence of this form of forgetting illustrates how retrieval processes shape memory not only by strengthening experiences that we recall regularly but also by impairing those memories whose recurring activation would impede retrieval.

The preceding view suggests that inhibiting related memories might be an adaptive response that reduces the disruptive effects of outdated information (Bjork, 1989). However, in many cases, the adaptiveness of inhibition seems less obvious. Consider recalling facts about a canary. On the preceding view, recalling one property of canaries should suppress other properties activated during recall, for example, recalling *canaries have wings* should suppress *canaries can fly*. Far from being outdated or disruptive, these related properties are central to representing what we know about the concept of interest. Our ability to reason about objects and events depends critically on our capacity to form and retain mental representations with many interrelated parts. If we are to avoid decomposing such representations, retrieval of any one part should proceed without suppressing the others: Recalling that cups have handles should not make us forget that cups can contain liquid, nor should recalling that presents are given at a birthday party make it surprising to hear about a cake. Indeed, people often retain stable representations of objects, particularly when their parts are interrelated. We are thus left with the question: How can retrieval cause forgetting of related memories and also allow for retention of the complex knowledge structures necessary to support cognition?

In this article, we explore an important boundary condi-

Michael C. Anderson and Kathleen C. McCulloch, Department of Psychology, University of Oregon.

The research reported herein was supported in part by startup funds from the University of Oregon. We thank Ted Bell, Robert Bjork, Dirk de Kok, Chad Dodson, Dan Kimball, Geeta Shivde, Bobbie Spellman, and the members of cog fog for comments on an early draft of this article. We also thank Ted Bell, Dirk de Kok, Emily Holzberg, and Toni Timmers for assistance with data collection and analysis.

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.

tion on retrieval-induced forgetting that suggests a partial answer to this question. The boundary condition is based on the old idea that retrieval competition may be reduced by forming interconnections between competing memories—a process called *integration* (cf. Smith, Adams, & Schorr, 1978). The idea that integration might limit interference emerged in the classical interference era, during which it was found that proactive and retroactive interference effects are dramatically reduced by interrelationships between competing responses (for reviews, see Horton & Kjeldergaard, 1961; Jenkins, 1963; Kjeldergaard, 1968; Osgood, 1946, 1949; Postman, 1971). More recently, studies have shown that interrelationships between facts about a concept eliminate the competition that often impedes retrieval of those facts (Moeser, 1977, 1979a, 1979b; Myers, O'Brien, Balota, & Toyofuku, 1984; Radvansky, Spieler, & Zacks, 1993; Radvansky & Zacks, 1991; Smith, Adams, & Schorr, 1978). Integration has even been proposed to explain how experts, who know many related facts, nonetheless seem less subject to interference (Smith et al., 1978). In this article, we explore whether integration may also protect memories from retrieval-induced forgetting. We believe that exploring this question will help solve the puzzle of how retrieval can play such a powerful role in shaping long-term accessibility while allowing the construction and use of representations that are resistant to inhibition.

Two bodies of research suggest that integration might protect people from retrieval-induced forgetting. First, work on fact recognition has shown that when several facts about a concept are interrelated, the competition that would occur during the retrieval of one of those facts is circumvented. In these studies, retrieval competition usually is measured as an increase in reaction time or error rate arising when people must recognize a previously studied fact (e.g., *A yuppie is in the park*) that uses concepts also contained in other studied facts (e.g., *A yuppie is in the church, A yuppie is in the store*). Although fact verification speed usually gets worse as more facts are learned about a concept (a finding known as the fan effect—see Anderson, 1974), this competition is often eliminated when those facts can be integrated. For instance, Radvansky and Zacks (1991) found that when subjects learned facts about a single location containing many objects (e.g., facts about a lobby—*the potted palm is in the lobby*), there was virtually no cost in verification time when facts about new objects (*the desk is in the lobby, the ashtray is in the lobby*) were learned. Little interference was found, despite an increase in the number of facts associated to the location cue. However when subjects learned facts about an object that could be in many locations (e.g., facts about a potted palm—*the potted palm is in the lobby, the potted palm is in the park*), there was a sizable cost in verification time when facts about new locations (e.g., *the potted palm is in the office*) were learned, replicating the usual fan effect. Radvansky and Zacks argued that when one can integrate facts about a concept (e.g., a location concept, such as *lobby*) into a coherent situation model, the facts no longer compete during retrieval because they are part of a larger mental representation.

Similar results to Radvansky and Zacks's (1991) study have been found in other fact retrieval studies in which integration was done in many different ways. Smith et al. (1978) found that learning additional facts about a person eliminated interference if those facts were consistent with a preexisting script. For instance, the new fact *Marty was chosen to christen the ship* eliminated fan interference between other facts, such as *Marty broke the bottle* and *Marty did not delay the trip.* The introduction of a theme (e.g., ship christening) helps consolidate these disparate facts. In a related finding, Myers, O'Brien, Balota, and Toyofuku (1984) found that when facts about an individual were related by causal or explanatory links (e.g., *The teacher read the menu, The teacher found that the prices were high, The teacher checked his wallet*), many new facts could be learned with no cost in retrieval time. Despite the variety of ways in which integration has been manipulated (McCloskey & Bigler, 1980; Moeser, 1977, 1979a, 1979b) all of these studies suggest that integrating facts associated with a concept reduces competition.

Integration benefits are not limited to speeded fact recognition. Findings from the classical interference era show that integration also reduces both proactive and retroactive interference. Studies from this period examined the usefulness of "mediate associates" in reducing interference (for reviews, see Horton & Kjeldergaard, 1961; Jenkins, 1963; Kjeldergaard, 1968; Postman, 1971). For instance, consider a classic study on verbal mediation by Russell and Storms (1955). In this study, subjects learned a first list of paired associates (each made of a nonsense syllable and a noun—e.g., *cef–stem*), followed by a second list in which each stimulus (e.g., *cef*) was paired with a new response. Russell and Storms found that the presence of a relationship mediating the responses on List 1 (e.g., *cef–stem*) and List 2 (e.g., *cef–smell,* which is related to *stem* via their mediate associate *flower*), dramatically reduced proactive interference, relative to a condition in which List 2 and List 1 responses were unrelated (e.g., *cef–stem* and *cef–joy*). Later work by Martin and Dean (1964) showed that these benefits depended on whether subjects deliberately used such associations to link responses: Most subjects reported using List 1 responses to help them recall List 2 competitors, and such claims predicted reduced interference. Martin and Dean found an even larger decrease in proactive interference when they asked subjects to use mediation. In a related study, Postman found that subjects spontaneously adopted a mediational strategy after practice with several paired-associate lists, even when competitors were unrelated (Postman, 1964). The reduced interference found in these studies was also seen in work on retroactive interference (Horton & Wiley, 1967; Kanungo, 1967; Schulz, Liston, & Weaver, 1968; Young, 1955), suggesting that integration facilitated both learning new competitors and retaining existing ones. Thus, the notion that subjects can use integration to circumvent interference seems well supported, at least when studied in classical paired-associate interference paradigms.

The preceding review shows that integration is a powerful means of reducing retrieval interference. It thus seems possible that integration might form an important boundary condition on retrieval-induced forgetting. There are two reasons why the extension of integration benefits should not be presumed, however. First, the mechanisms underlying retrieval-induced forgetting might differ from those at work in the paradigms usually used to study integration benefits. For example, repeatedly recalling several members of a studied category (e.g., *Fruits*), impairs subjects' long-term retention of high-frequency (e.g., *Banana*) but not low-frequency exemplars (e.g., *Guava*), as would be expected if the more likely to interfere high-frequency exemplars required suppression more often than the low-frequency exemplars (Anderson et al., 1994). Repeated retrieval also causes generalized suppression that is not easily explained by traditional associative interference processes (Anderson & Bjork, 1994). For instance, retrieving some category members (e.g., *Red Blood*) has been shown to impair competing exemplars, regardless of whether they are tested with the same cue used to do repeated retrieval (e.g., Red *Tomato*) or a different cue (e.g., Food *Cherry;* Anderson & Spellman, 1995). Because we do not know how integration might alter the impact of these inhibitory processes, and because the role of such processes in prior integration studies is unclear, the extension of integration benefits to retrieval-induced forgetting should be studied directly.

A second reason to directly assess this extension of integration benefits stems from novel features of the behavioral situation addressed by retrieval-induced forgetting. Few studies have examined whether integration benefits would endure if a subset of the integrated associates was repeatedly and selectively retrieved. Research on part-set cuing impairment casts doubt on whether integration benefits would persist in this situation. In the part-set cuing inhibition procedure, subjects study several members of a common "set" for a later test (e.g., *Dove, Sparrow, Robin,* and *Canary* as members of *Birds*). Later, it is often found that providing members of the set (e.g., *Dove* and *Robin*) as cues actually impairs recall of remaining items (e.g., *Sparrow* and *Canary;* Slamecka, 1968; see also Mueller & Watkins, 1977; Nickerson, 1984; Roediger, 1973; Rundus, 1973). Recently, Basden and Basden (1995) suggested that this impairment might arise because part-set cues disrupt the use of recall strategies that subjects developed while studying. In this idea, when subjects study a set of words (e.g., *Birds*) they form a plan for recalling items that conforms to how they have structured them in memory. When the experimenter provides part of the study set as cues, however, subjects deviate from their initial plan and use whatever recall strategy is suggested by the cues. If the cues suggest a strategy that is incongruent with the subject's initial organization, recall of remaining items may be worse than recall of those same items when no cues are provided. Basden and Basden (1995) explored this *strategy-disruption hypothesis* in several studies, providing part set cues that were either consistent or inconsistent with how subjects organized items during study. In general, they found that cue sets that conflicted with subjects' initial organization caused more impairment, as predicted by the strategy disruption hypothesis.

The findings of Basden and Basden (1995) suggest a way that integration may actually increase retrieval-induced forgetting. If presenting a portion of an integrated set on a final test can cause reorganization that disrupts subjects' recall strategies, perhaps repeatedly retrieving a portion of an integrated set might have similar effects. Consider a subject who has studied four items—Ax, Saw, Glass, and Teeth—under a shared cue. If instructed to integrate these items, subjects might form a retrieval strategy that links items by a common feature—say "sharp." On a later test, this strategy may improve recall over what it would have been had no interrelationships been formed. However, if a subset of these items was repeatedly retrieved—for instance, Ax and Saw—before the final test, the initial strategy may be less useful, particularly if the retrieved set highlights another feature not shared with remaining items (e.g., Tools). Thus, subjects' recall of previously practiced items (e.g., Ax and Saw) on the final test may divert them from their initial strategy (e.g., Sharp). Compared with a condition in which the subset was not retrieved, recall for remaining items may be worse. Thus, even if integration reduces competition, it may not protect people from retrieval-induced forgetting—indeed, it may increase it.

In this article, we report three studies that assess the impact of integration on retrieval-induced forgetting. We adapted the retrieval-practice procedure of Anderson et al. (1994) to create varying levels of integration. The standard version of this procedure has four phases: a study phase, a retrieval-practice phase, a distractor phase, and a final test phase. In the study phase, subjects study six members from each of eight categories (e.g., fruit), with each item presented in category-exemplar format (e.g., *Fruit Orange*) on separate pages of a study book. Next, they perform directed retrieval practice by completing category-stem cued recall tests (e.g., *Fruit Or__*) on half of the exemplars from half of the studied categories; each item is practiced three times to increase the impact of these practices on the later recall of competing items (e.g., *Fruit Banana*). Following a 20-min retention interval, subjects are cued with each category name (e.g., *Fruits, Drinks*) and asked to recall all members of that category that they remember having seen in the experiment. Recall performance is measured on three types of item: (a) practiced items from practiced categories (hereafter called Rp+ items, e.g., *orange*); (b) unpracticed items from practiced categories (Rp− items, e.g., *Banana*); and (c) unpracticed items from unpracticed study categories (Nrp items, e.g., *Scotch*). On the final test, it is usually found that having done retrieval practice on some members of a category (Rp+ items) impairs recall for the remaining nonpracticed items (Rp− items) in those categories relative to performance for baseline study items (Nrp items; Anderson et al., 1994; Anderson & Spellman, 1995).

To see whether integration might affect retrieval-induced forgetting, we manipulated the instructions given during the study phase of the retrieval-practice procedure. In the standard instructions, subjects are asked to study each pair by relating the exemplar to its category label. This instruction encourages subjects to consider the category-exemplar association in isolation, although it does not preclude the possibility that they will interrelate exemplars with one another. All experiments in this article include a replication of this standard condition. A second group received instructions that were identical, except for the addition of one or more statements encouraging subjects to form relationships between exemplars. In this article, we refer to the use of these modified instructions as an integrative-rehearsal condition. If explicitly encouraging subjects to integrate items increases the amount of integration above what occurs spontaneously in the standard condition, we can assess the impact of integration on retrieval-induced forgetting.

Of course, if the degree of *spontaneous integration* is high in the standard condition, we may not see differences in retrieval-induced forgetting across these conditions. Thus, our second aim was to assess whether subjects in the standard condition spontaneously used strategies fostering integration, and whether such spontaneous integration has the same effect as our instructions. To find out, we included a postexperimental questionnaire designed to measure spontaneous rehearsal and integration done during the study phase. Although such retrospective reports should be interpreted with caution, Martin and Dean's (1964) findings suggest that subjects' introspections about integration were effective at predicting proactive interference in the Russell and Storms (1955) procedure. Following on Martin and Dean's example, we used subjects' questionnaire responses to divide the standard condition into *high*- and *low-integrator* groups, to see whether these subjective differences might parallel the findings of our explicit manipulation. If spontaneous integration leads subjects to integrate exemplars, we would expect the retrieval-induced forgetting observed for high integrators in the standard condition to be similar in size to that seen for the integrative rehearsal condition. On the other hand, the amount of inhibition observed for low integrators should differ from that found for both high integrators and subjects in the integrative-rehearsal condition.

Experiment 1

In Experiment 1, we manipulated the instructions given during the study phase of the retrieval-practice procedure. The standard group was simply asked to study each pair by forming a relationship between the category label and the example (e.g., *Fruit Orange*). In the integrative-rehearsal condition, we also asked subjects to rehearse each item with previously studied items from that category. After this variation in study phase instructions, subjects were treated identically. They performed retrieval practice on four of the eight categories and, after a 20-min delay, received a category cued recall test for all studied categories. At the end, standard subjects completed an integration questionnaire.

If retrieval-induced forgetting is found, doing retrieval practice on Rp+ items should impair final recall performance for Rp−, relative to baseline Nrp items. If integration reduces retrieval-induced forgetting, Rp− items should be less impaired in the integrative-rehearsal than in the standard condition; furthermore, if spontaneous integration effects parallel those from our instructional manipulation, the difference in Rp− impairment should be largest between the integrative-rehearsal and standard-low groups.

Method

Subjects

One hundred forty-four University of Oregon undergraduates participated to fulfill a course requirement. Subjects were randomly assigned to the integrative-rehearsal ($n = 48$) or to the standard condition ($n = 96$).

Design

Two variables, study instructions and retrieval-practice status, were manipulated. Study instructions was manipulated between subjects and had two levels: standard and integrative-rehearsal. Standard subjects were also divided into two groups, the low and high integrator groups, based on their responses to a postexperimental questionnaire (see *Results and Discussion* section), yielding three groups in total: the standard-low, standard-high, and integrative-rehearsal groups.

The retrieval-practice status of an item was manipulated within-subjects. Exemplars either (a) were given retrieval practice (Rp+ items); (b) were unpracticed but were members of a practiced category (Rp− items); or (c) were unpracticed and members of an unpracticed study category (Nrp items). The dependent measure was the proportion of items correctly recalled in each condition on a category-cued recall test.

Materials

Category construction. Ten categories, two of which were fillers, were drawn from Battig and Montague (1969) and McEvoy and Nelson (1982). Six exemplars were chosen for each category. Exemplars were moderate to high in taxonomic frequency (average position = 7.2 in Battig & Montague, 1969) and low in overall word frequency ($M = 6$ per million in Kučera & Francis, 1967). Within each category, exemplars began with distinct letters. Every studied item also began with a distinct two-letter stem, ensuring that stems were unique for retrieval practice. Stem difficulty was controlled by using stems with high versatility (i.e., the number of words in Kučera and Francis starting with that stem—Solso & Juel, 1980; $M = 242$). To reduce interference from nonstudied items, we tried to ensure that the first two letters of studied items did not cue nonstudied exemplars listed in Battig and Montague (1969). See the Appendix for these materials.

Learning books. Each page of the study book displayed a category with an exemplar (e.g., *Dogs Collie*). Booklets contained 60 items, 6 items from each of eight experimental and two filler categories. To match the study position of every category and condition, the presentation order was block randomized, yielding six blocks with one item from each category. To control for primacy and recency effects, the first and last two items in the list were fillers. Three versions were made by assigning items to blocks three times.

Retrieval-practice books. Each page of the retrieval-practice book displayed a category and the first two letters of an item (e.g., *Dogs Co___*). The 48 test pages were ordered according to several constraints. To control for primacy 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, 3.4 items intervened between the first and second practice test and 6.7 items between the second and third. No two items from a category were tested adjacently, and the mean test position of categories was matched ($M = 23.5$). When possible, we kept sequences of tests from repeating by adding tests of fillers.

To ensure that every item appeared in every condition, we counterbalanced which categories were practiced and also which exemplars within a category were practiced. We did this by dividing the eight categories into two sets of four (Sets A and B), and dividing exemplars within categories into two sets of three (see Anderson et al., 1994, for details). This process yielded four sets of 12 items (A1, A2, B1, and B2) which, together with the ordering constraints described previously, were used to make four books. The distractor phase booklet included reasoning problems with no words used in the main experiment.

Test books. A filler category appeared on the first page of the test book, to acquaint subjects with the procedure, with the following eight pages testing critical categories. To control output interference across Rp and Nrp categories, the mean test position of the A and B sets was matched. Based on these constraints, four books were made by rotating the four categories from each set through the test positions assigned to it. Across subjects, all categories were equated for test position.

Questionnaire. A response sheet included the question: "During the study phase, when you saw a word pair, how often, if ever, did you intentionally think back to previously seen category members? For example, if you first saw Cars Honda and then Cars Ford, did you intentionally rehearse Ford and Honda together?"

Procedure

Experiment 1 was conducted in three phases: a study, a retrieval practice, and a category-cued recall phase. In the study phase, subjects were randomly assigned to one of three study booklets. Subjects in both groups were told that they would see category-example pairs and that they should study each pair for 6 s by relating the example to its category name. Subjects in the integrative-rehearsal condition were also asked to rehearse each example with all the examples they could remember having studied up until that point.

In the retrieval-practice phase, subjects received one of four booklets, depending on the counterbalancing group (see the *Materials* section). Subjects were told that each page contained a category with the first two letters of a studied exemplar, which they were to recall and write down within 7 s, along with the category name. Subjects were warned that items would be tested several times, and that they should try to write the correct item each time. After the practice phase, subjects did a reasoning task for the 20-min retention interval.

For the final test, subjects got one of the four booklets. Subjects were told that each page contained a category name, and that they would have 30 s to recall as many exemplars as they could remember from the first two phases. After the experiment, standard subjects completed an integration questionnaire.

Results and Discussion

All analyses were performed with group (standard-low, standard-high, and integrative-rehearsal) and practice counterbalancing as between-subjects variables.[1] An alpha level of .05 was used for all statistical tests.

Assignment of Subjects in the Standard Condition to Low and High Integrator Groups

Standard subjects' questionnaire responses were rated on a scale ranging from 1 (*no integration*) to 4 (*extensive integration*) by 3 raters who came to 100% agreement on each subject. Subjects were rank ordered by rating within each practice counterbalancing condition, with the bottom halves of these orderings comprising the standard-low group (mean rating = 1.58) and the top halves the standard-high group ($M = 3.23$).

Retrieval Practice

Retrieval-practice success rates did not vary across the standard-low ($M = 86\%$), the standard-high ($M = 88\%$), and the integrative-rehearsal conditions ($M = 87\%$), $F < 1$.

Final Test Performance

We first examined the data collapsed over groups to verify that the procedure had the expected positive and negative effects on Rp+ and Rp− items, respectively. As predicted, retrieval practice facilitated recall of practiced items (Rp+ = 73%) relative to baseline items (Nrp = 45%) on the final test, $F(1, 132) = 391.95$, $MSE = 2.61$, whereas it impaired Rp− items ($M = 37\%$) relative to Nrp baseline items (45%), $F(1, 132) = 52.34$, $MSE = 1.81$. These findings replicate the retrieval-based learning and retrieval-induced forgetting effects observed by Anderson et al. (1994).

Variations in facilitation and impairment. Next, we examined facilitation and retrieval-induced forgetting in our three groups. As illustrated in Table 1, all groups showed substantial Rp+ facilitation, with similar amounts observed for the integrative-rehearsal and the standard-low groups, $F < 1$. Standard-high subjects showed numerically less facilitation, although the amount was not reliably smaller than that found for the standard-low, $F(1, 132) = 2.73$, $MSE = 2.61$, or for the integrative-rehearsal group, $F(1, 132) = 2.40$, $MSE = 2.61$. Thus, retrieval-based facilitation is comparable in these groups.

More important, however, are variations in retrieval-induced forgetting. If integration benefits extend to retrieval-induced forgetting, less impairment should be found for standard-high and integrative rehearsal subjects. Consistent with this view, standard-low subjects showed more impairment than did integrative-rehearsal subjects, $F(1, 132) = 5.54$, $MSE = 1.81$, and they also showed marginally greater impairment than did standard-high subjects, $F(1, 132) = 3.23$, $p = .07$, $MSE = 1.81$. Impairment for standard-high and integrative-rehearsal subjects did not differ, $F < 1$. Thus, our standard group was made of a mixture of people,

[1] Because the difference between the standard-low and standard-high groups was determined post hoc, any differences seen between these two sets of subjects or between either one of these sets and the integrative-rehearsal condition must be interpreted with appropriate concern for subject variables (other than reported integration) that may contribute to the patterns observed. To highlight this distinction, we adopt the convention of referring to subjects that receive different treatments as being in different conditions, and to subjects differing on the basis of spontaneous integration as being in different groups.

Table 1
Mean Percentage of Words Correctly Recalled as a Function of the Retrieval-Practice Status of an Item and Rehearsal Group in Experiment 1

	Retrieval-practice status of an item				
Rehearsal group	Rp+	Rp−	Nrp	Inhibition (Nrp − Rp−)	Facilitation (Rp+ − Nrp)
Standard-low	74	33	45	12	29
Standard-high	73	43	50	7	23
Integrative-rehearsal	71	37	43	6	28
M	73	37	45		

Note. Rp+ = items within a category that received retrieval practice; Rp− = unpracticed members of a practiced category; Nrp = unpracticed items in an unpracticed category; standard-low and standard-high = groups containing subjects who were not given rehearsal instructions during the initial study phase of the experiment, but who later reported either low or high rates of spontaneous integration. Integrative-rehearsal = subjects who were explicitly instructed to rehearse.

some of whom spontaneously rehearsed, showing reduced impairment like integrative-rehearsal subjects, and others of whom did not rehearse, showing more impairment. Despite these reductions, reliable impairment was found in the standard-low, $F(1, 132) = 37.63$, $MSE = 1.81$, the standard-high, $F(1, 132) = 12.91$, $MSE = 1.81$, and in the integrative-rehearsal condition, $F(1, 128) = 7.86$, $MSE = 1.81$. Thus, neither subjects' spontaneous strategies nor our instructions were sufficient to eliminate retrieval-induced forgetting. However, these findings clearly support the idea that integration protects the associates of a cue from the inhibitory effects of retrieval.

Potential concerns. Despite these positive findings, several concerns should be resolved before concluding that integration reduces retrieval-induced forgetting. The first has to do with differences in the strength of category-exemplar associations across standard and integrative-rehearsal subjects. Subjects in both conditions had 6 s to study each pair. However, standard subjects only had to study the category-exemplar relationship, whereas integrative-rehearsal subjects had to do this and rehearse items with one another. Given the limited time, these subjects may not have learned category-exemplar associations as effectively as the standard subjects did. If weakly associated exemplars interfered less during retrieval practice of their category mates, they may not have been suppressed. Thus, retrieval-induced forgetting might have been reduced even if integration had no beneficial effect. To the extent that integration is confounded with this difference in representation, we cannot be sure that it reduces retrieval-induced forgetting. Thus, the present findings need to be replicated in a way that controls for potential differences in the strength of Rp− category-exemplar associations.

Finally, although our instructional manipulation reduced impairment, we might have manipulated integration more strongly. Rehearsing exemplars could be a weak way of achieving integration. Research on fan effects emphasized the formation of meaningful relations between competitors as central to the benefits of integration (Smith et al., 1978). Impairment might be reduced more completely if our instructions explicitly emphasized the formation of such relations. Experiment 2 addresses these questions.

Experiment 2

In Experiment 2, our aim was to replicate the reduced retrieval-induced forgetting found in Experiment 1 and to see if stronger integration instructions might reduce impairment further. To strengthen our manipulation, we made the goal of integration explicit by asking subjects to interrelate items in a meaningful way. No method was given for finding interrelationships, because we wanted subjects to use whatever strategy was most natural. If our instructions foster integration, and if integration reduces impairment, integrative-rehearsal subjects should show less impairment than standard subjects.

In this study, we also controlled the amount of time subjects had to encode category-exemplar associations. In Experiment 1, both standard and integrative-rehearsal subjects had 6 s of study time, even though the latter group had to both encode the category-exemplar relation and rehearse items together. To rectify this inequity, we gave integrative-rehearsal subjects two passes through the study book. In the first pass, both conditions got the same study instructions, focusing on the formation of category-exemplar relationships. Integrative-rehearsal subjects then got a second pass through the book, during which they had 7 s to interrelate each item with other items in its category. This procedure ensured that integrative-rehearsal and standard subjects had the same initial time to form category-exemplar associations. If integrative-rehearsal subjects show less impairment, it cannot arise from deficits in encoding category-exemplar relationships.

Naturally, adding a study phase introduces differences other than our integration instructions. Our second study phase both adds a repetition and increases study time by over a factor of 2. Although prior work (Anderson et al., 1994) suggests that increased study time might work against finding integration benefits (by increasing Rp− interference), the effects of this variable on retrieval-induced forgetting have never been documented. Because of this uncertainty, we must establish that increased study time alone does not reduce retrieval-induced forgetting.[2] To explore the effects of study time on retrieval-induced forgetting, we added a control group who also received two study passes, but who were simply told they had an extra 7 s to study. To distinguish it from the primary control condition, which was also given standard instructions, we refer to this control as the standard-twice condition (i.e., standard subjects, two study passes) and to the primary control as the standard-once condition. If extra study reduces retrieval-induced forgetting, standard-twice should be less impaired than standard-once subjects, even though no integration instructions were given. Furthermore, if extra time alone reduced impairment for integrative-rehearsal subjects, impairment should not vary across the standard-twice and integrative rehearsal conditions.

Reduced impairment in the standard-twice condition would not imply that adding study time (or repetitions) per se reduces retrieval-induced forgetting, however. Adding study time is likely to increase spontaneous integration (Bousfield, 1953; Tulving, 1962). Such an increase is likely to reduce impairment. To test this idea, we again included a final questionnaire to measure subjects' integration tendencies. To make responses more quantifiable, we replaced the verbal report method of Experiment 1 with a scale ranging from 1 (*none of the time*) to 5 (*all of the time*). If our scale accurately measures subjects' integration, higher ratings should be given by integrative-rehearsal than standard-once subjects. Furthermore, if adding study time increases spontaneous integration, ratings should be higher for standard-twice than for standard-once subjects.

A final goal was to replicate the reduced retrieval-induced forgetting found for high spontaneous integrators in Experiment 1. In this study, we also divided subjects into low and high integrators based on questionnaire responses. If spontaneous integration reduces retrieval-induced forgetting in the same way as our instructional manipulation, high integrators should be less impaired than low integrators. Furthermore, high integrators may show more benefit in the standard-twice condition, because they have more time to integrate. To exaggerate this advantage, we reduced study time for the first study pass (for all conditions) to 4 s a pair, making it harder for high integrators to integrate in the standard-once condition. Finally, because integrative-rehearsal subjects were specifically asked to integrate, differences between the high and low groups should be diminished.

Method

Subjects

Ninety-six University of Oregon undergraduates participated to fulfill a course requirement. Subjects were randomly assigned to the standard-once ($n = 32$), the standard-twice ($n = 32$), or the integrative rehearsal ($n = 32$) condition.

Design

Study condition was manipulated between-subjects, on three levels: the standard-once, standard-twice, and integrative-rehearsal conditions. As in Experiment 1, retrieval-practice status was manipulated within subjects, and questionnaire responses were used to construct low- and high-integration groups within each condition.

Materials

The materials of Experiment 1 were used, except for the following changes.

Study and retrieval-practice lists. Because standard-twice and integrative rehearsal subjects viewed the study list twice and were asked to write tick marks on the study pages (see the *Procedure* section) during the second viewing, two copies of the list were made for each subject. Only one of the three study orders developed for Experiment 1 was used.

Questionnaire. The questionnaire asked subjects to estimate how much time they spent interrelating exemplars during the study phase. Although this question resembled that of Experiment 1, subjects responded on a 5-point scale for each category. Next to each category was a scale ranging from 1 to 5, with the numbers 1, 3, and 5 labeled with *none of the time, some of the time,* and *all of the time,* respectively. Subjects were encouraged to use the entire scale. Finally, we emphasized that there were no right or wrong answers and that they should put ratings that truly reflected their experience.

Procedure

The procedure of Experiment 1 was used, except for the following changes.

First study phase. We changed the presentation rate of the initial study phase from 6 to 4 s per pair to reduce spontaneous integration. Following this phase, standard-once subjects did retrieval-practice whereas the remaining subjects got a second study attempt.

Second study phase. Standard-twice and integrative rehearsal subjects were told that they were being given an extra 7 s to study each pair. Integrative-rehearsal subjects were also asked to interrelate each exemplar with as many other exemplars of that category that they remembered studying up to that point. To keep subjects oriented to the integration task, they were asked to indicate how many studied exemplars they were able to relate to the current one by making tick marks under the current pair for each additional item they integrated. The presentation order of pairs was not varied between the two study phases.

Questionnaire. Because of the simplified format (see *Materials* section), subjects were given 2 instead of 5 min to respond to the questionnaire.

[2] Extra study time might reduce impairment if additional study mainly affected the representation of the exemplar rather than the category-exemplar association, for example. Strengthening the target representation could reduce the impact of inhibition, even if strengthening the cue-target representation made the item more interfering. Because the effects of these factors have never been empirically separated, it is prudent to ensure that study time by itself does not reduce inhibition, despite expectations to the contrary.

Table 2
Mean Percentage of Words Correctly Recalled as a Function of the Retrieval-Practice Status of an Item and Rehearsal Condition in Experiment 2

	Retrieval-practice status of an item				
Rehearsal group	Rp+	Rp−	Nrp	Inhibition (Nrp − Rp−)	Facilitation (Rp+ − Nrp)
Standard-once					
Low integrators	73	26	38	12	35
High integrators	80	31	40	9	40
M	77	29	39	10	38
Standard-twice					
Low integrators	81	50	60	10	21
High integrators	78	54	56	2	22
M	80	52	58	6	22
Integrative-rehearsal					
Low integrators	71	44	49	5	22
High integrators	72	55	55	0	17
M	72	49	52	3	20

Note. Rp+ = items given retrieval practice; Rp− = unpracticed items in a practiced category; Nrp = baseline items; standard-once = subjects given one pass through the study list and no special study instructions; standard-twice = subjects given two passes through the study list and no special instructions; integrative-rehearsal = subjects given integrative-rehearsal instructions on the second pass through the study list; low and high integrators = subjects reporting low versus high integration on the questionnaire.

Results and Discussion

All analyses included study instructions, practice counterbalancing, and reported integration (as measured by our questionnaire) as between-group factors. Practice counterbalancing did not interact with the other factors and will not be discussed further.

Assignment of Subjects to Low and High Spontaneous Integration Groups

We made a rehearsal score for each subject by averaging their rehearsal ratings across the eight categories. Using this score, we followed the procedure of Experiment 1, to make the low (M score = 2.98) and high groups (M = 3.72). This difference in score was similar in the standard-once (M = 2.52 vs. 3.23, for low vs. high), the standard-twice (M = 3.12 vs. 3.94), and the integrative-rehearsal conditions (M = 3.31 vs. 4.00).

Retrieval-Practice Performance

Success rate was similar in the standard-once (M = 90%), standard-twice (M = 92%), and integrative-rehearsal conditions (M = 90%), $F < 1$, and in the low (M = 91%) and high integrator groups (M = 91%), $F < 1$.

Final Recall Performance Averaged Across Subjective Integration Reports

In presenting the final recall data, we first consider our manipulation of Study Instructions; we then discuss the effects of spontaneous integration. Table 2 shows the recall percentages for each item type in the three study conditions. As expected, studying items twice improved recall, with standard-once subjects recalling fewer items (M = 46%) than standard-twice (M = 62%), $F(1, 72) = 24.50$, MSE = 9.81, and integrative-rehearsal subjects (M = 57%), $F(1, 72) = 10.72$, MSE = 9.81. These findings also replicate the overall facilitation and impairment of Rp+ and Rp− items, respectively: Retrieval practice improved the final recall of Rp+ (M = 76%) relative to Nrp baseline items (M = 50%), $F(1, 72) = 328.87$, MSE = 1.95, whereas it impaired recall of Rp− (M = 43%) relative to Nrp items (50%), $F(1, 72) = 23.79$, MSE = 1.58.

Variations in retrieval-induced forgetting. More crucial is whether our new instructions reduced retrieval-induced forgetting. As can be seen in Table 2, our new integration instructions reduced impairment, relative to that found for the standard-once condition, $F(1, 72) = 5.37$, MSE = 1.58. This reduction replicates that of Experiment 1,[3] and shows that integration benefits cannot be explained solely by diminished encoding of Rp− category-exemplar associations: integrative-rehearsal subjects had as much, if not more, time for this type of encoding as standard-once subjects. Indeed, impairment was reliable for standard-once, $F(1, 72) = 20.61$, MSE = 1.58, but not for integrative-

[3] The present result demonstrates that integrative-rehearsal subjects show less impairment than uninstructed-once subjects on an absolute scale. One might be concerned that absolute differences like this are hard to interpret in light of the differences in baseline Nrp levels used to measure impairment. However, when impairment is expressed as a proportional loss, relative to the Nrp baseline, differences in retrieval-induced forgetting between the integrative-rehearsal and standard conditions actually grow. This arises because the condition with the smaller absolute impairment (integrative-rehearsal condition) also has the higher baseline. Nevertheless, any doubts that might arise from these differences in baseline are addressed in Experiment 3.

rehearsal subjects, $F(1, 72) = 1.60$, $MSE = 1.58$. Thus, our new instructions enhanced our manipulation of integration.

In contrast to integrative-rehearsal subjects, the standard-twice condition showed retrieval-induced forgetting, $F(1, 72) = 6.99$, $MSE = 1.58$, indicating that two study exposures does not by itself eliminate impairment. Yet, standard-twice impairment did not differ reliably from that found for standard-once, $F(1, 72) = 1.80$, $MSE = 1.58$, or integrative-rehearsal subjects, $F < 1$. Thus, studying items twice may also reduce impairment. We discuss why this reduction may have occurred in our section titled *Characteristics of questionnaire responses.*

Differences in facilitation. Next, we examined whether Rp+ facilitation varied across conditions. As can be seen in Table 2, practiced items in the standard-once condition showed more facilitation over baseline than did those items in the standard-twice, $F(1, 72) = 20.92$, $MSE = 1.95$, or integrative-rehearsal conditions, $F(1, 72) = 26.62$, $MSE = 1.95$. This difference was substantial and unexpected. It may reflect ceiling effects: Perhaps Rp+ recall is already so high in the standard-once condition that further increments are difficult to observe. Alternatively, less facilitation may reflect the diminishing returns typical of the power law of practice (Newell & Rosenbloom, 1981). If genuine, this difference makes it more difficult to interpret reduced impairment for the integrative-rehearsal condition. We return to this issue in the our section titled *Alternative Interpretations of Integration Effects.*

Final Recall Performance Broken Out by Subjective Integration Reports

Characteristics of questionnaire responses. Our new integration scale was sensitive to differences in integration across conditions. Integrative-rehearsal subjects reported more integration ($M = 3.65$) than did standard-once subjects ($M = 2.87$), $F(1, 72) = 24.80$, $MSE = .39$, who reported less than standard-twice subjects ($M = 3.53$), $F(1, 72) = 17.50$, $MSE = .39$. This latter finding supports our prediction that spontaneous integration would increase with more study time. Although less integration was reported in the standard-twice than in the integrative rehearsal condition, this difference was not reliable, $F < 1$. Thus, reduced impairment in the standard-twice condition may stem from increased spontaneous integration.

Differences in recall across low and high integrators. Table 2 shows the recall percentages for each item-type for low and high integrators, within our three instructional conditions. Low and high integrators were similar in overall recall ($M = 53\%$ and 56% for low and high), $F < 1$, and in the degree of Rp+ facilitation (for low, Rp+ − Nrp = 75 − 49 = 26%; for high, 77 − 51 = 26%), $F < 1$. Despite these similarities, low integrators showed marginally more impairment (Nrp − Rp− = 49 − 40 = 9%) than high integrators (51 − 47 = 4%), $F(1, 72) = 3.58$, $p = .06$, $MSE = 1.58$, replicating the questionnaire results of Experiment 1. These findings also alleviate concerns that reduced retrieval-induced forgetting stems from diminished Rp+ facilitation. As in Experiment 1, reliable impairment was found for both low, $F(1, 72) = 22.91$, $MSE = 1.58$, and high integrators, $F(1, 72) = 4.46$, $MSE = 1.58$.

Next, we examined variations in how well integration reports predicted impairment across conditions. We speculated that such reports would be most diagnostic in the standard-twice condition because the added study time would exaggerate individual differences in integration. As seen in Table 3, standard-twice subjects did show the greatest difference in impairment (10% vs. 2% for low and high integrators), followed by integrative-rehearsal (5% vs. 0% for low and high) and standard-once subjects (12% and 9%).[4] However, these differences did not vary reliably across the conditions ($F < 1$ in all cases). Nevertheless, the overall difference in impairment between low and high groups argues that self-reports predict variations in retrieval-induced forgetting. As illustrated in Table 3, the pattern of impairment for low and high integrators nicely corresponds to that seen in our instructional manipulation. Indeed, the degree of reported integration (in parentheses) predicts the amount of impairment, irrespective of study time and instructions, $r = -.949$.

Alternative Interpretations of Integration Effects

The current findings replicate Experiment 1 and strongly support a role for integration in protecting against the inhibitory effects of retrieval. As Table 3 makes clear, our new instructions to interrelate items nearly eliminated impairment. These findings are paralleled by the data from subjects' retrospective integration reports, which also replicate results found in Experiment 1. Furthermore, the near absence of impairment for integrative-rehearsal subjects argues that diminished attention to category-exemplar associations during study cannot be the sole cause of integration benefits. Integrative-rehearsal subjects in Experiment 2 spent as much or more time encoding associations than did standard-once subjects, yet still showed reduced impairment. These results argue that impairment is reduced specifically because exemplars become integrated.

Despite these supportive results, several alternative hypotheses should be addressed before concluding that integration reduces retrieval-induced forgetting. In this section, we discuss two factors that may contribute to these benefits: Differential exposure time and differential competition. These alternatives motivate the manipulations in our final study.

Differential exposure time. Although integrative rehearsal subjects showed less retrieval-induced forgetting, they also had more study time than standard-once subjects.

[4] In some respects, the fact that low and high integrators in the integrative-rehearsal condition show the same pattern of impairment as in the other conditions is surprising. For instance, an instruction to integrate might have made people's self-reports of integration sensitive to different subject selection variables (e.g., tendency to comply with instructions) than were present in the standard conditions. We believe that this consistency arises because subjects, on the whole, try to comply with instructions and only vary in the effectiveness with which they comply—a dimension likely to correlate with a spontaneous disposition to integrate.

Table 3
Mean Percentage of Impairment of Rp− Items (and Mean Integration Rating) as a Function of the Manipulated Rehearsal Condition and the Degree of Integration Reported on the Questionnaire in Experiment 2

	Rehearsal condition		
Questionnaire integration group	Standard-once	Standard-twice	Integrative-rehearsal
Low integrators			
Impairment	12%	10%	5%
Mean rating	2.52	3.12	3.31
High integrators			
Impairment	9%	2%	0%
Mean rating	3.23	3.94	4.00

Note. Standard-once = subjects given one pass through the study list and no special study instructions; standard-twice = subjects given two passes through the study list and no special instructions; integrative-rehearsal = subjects given integrative-rehearsal instructions on the second pass through the study list; low and high integrators = subjects reporting low vs. high amounts of integration on the postexperimental questionnaire.

This factor may have reduced impairment. Consistent with this view, numerically less impairment was found in the standard-twice than in the standard-once condition; though significant, the impairment for standard-twice subjects was not reliably greater than that found for integrative-rehearsal subjects. Thus, the standard-twice condition may support a role for increased study time in reducing impairment. We have argued that these study time effects only arose because people had more time to spontaneously rehearse. Indeed, standard-twice subjects reported higher integration rates than standard-once subjects did, and these rates predicted impairment. Nevertheless, one might argue that increased study time may have contributed to reduced impairment in the standard-twice and integrative-rehearsal conditions. If so, it is desirable to show reduced impairment in the integrative-rehearsal condition above that found with increased study time.

Differential competition. The reduced impairment found for integrative-rehearsal subjects may stem from reduced Rp+ facilitation in that condition (M = 20%), relative to the standard-once condition (M = 38%). According to this hypothesis, strengthening Rp+ items impairs Rp− performance because Rp+ items exert more competition during the final test (Raaijmakers & Shiffrin, 1981; Rundus, 1973). If the degree of Rp+ strengthening determines impairment, deficient strengthening in the integrative-rehearsal condition may have reduced retrieval-induced forgetting. Thus, less impairment may have arisen even if integration had no effect. We refer to this as the differential competition hypothesis.

There are at least three reasons to believe that the present effects are not caused by differential competition. First, although integration benefits are confounded with reduced strengthening when we compare the integrative-rehearsal and standard-once conditions, this confounding does not arise in other aspects of the data that support integration effects. For example, in our questionnaire analysis, high integrators showed reduced impairment, even though strengthening for the low (26%) and high (26%) groups was similar. Second, prior work with the retrieval-practice paradigm has shown the degree of Rp+ strengthening to be a poor predictor of retrieval-induced forgetting (Anderson et al., 1994; Anderson, Bjork, & Bjork, 1993; Ciranni & Shimamura, 1998). Indeed, significant Rp+ strengthening can yield little impairment, provided that Rp− items are less likely to interfere during retrieval practice (Anderson et al., 1993, 1994). If strengthening was a poor predictor in prior studies, it is unclear why it should be better in this study.

Finally, reduced facilitation in the integrative-rehearsal condition may not imply less strengthening. Reduced facilitation may simply reflect ceiling effects on Rp+ performance or, perhaps, some factor affecting the expression of strengthening on the final test. For instance, integration may alter the output order of items on the final test (Bousfield, 1953; Tulving, 1962). If subjects integrated practiced and unpracticed items, they may be more likely to interleave Rp+ and Rp− items on the final test, because stronger Rp+ items would cue associated Rp− items. Such a recall pattern would differ from the typical case (Anderson et al., 1994) in which Rp+ items, because of their greater strength, are recalled, on average, before Rp− items. To the extent that integration leads subjects to interleave Rp+ and Rp− items, reduced facilitation may reflect greater output interference on practiced items, relative to the standard-once condition. Spontaneous integration in the standard-twice condition might reduce facilitation for similar reasons.

If integration altered the output order of Rp+ and Rp− items, we should see less clustering of Rp+ items (i.e., more interleaving of Rp+ with Rp− items) on the final test. To explore this idea, we applied a standard measurement of clustering—the adjusted ratio of clustering (hereafter, ARC; Roenker, Thompson, & Brown, 1971)—to the recall order of items within each category, treating Rp+ and Rp− items as two "sets" that might or might not cluster. For each subject, we derived a metric of Rp+/Rp− clustering by computing an ARC score for each practiced category and averaging those scores over categories. The ARC value reflects the degree to which clustering exceeds that expected by chance, with 0 indicating chance clustering, +1 indicating perfect clustering, and negative scores indicating clustering below chance. On the basis of Anderson et al. (1994), we should find positive ARC scores for practiced categories in the standard-once condition, reflecting subjects' tendency to recall Rp+ items early and together. However, if integration increases the interleaving of Rp+ and Rp− items, clustering should be reduced in the integrative-rehearsal, relative to the standard-once condition.

As expected, standard-once subjects showed more clustering (ARC = .42) than did subjects in the integrative-rehearsal (ARC = −.04), $F(1, 72) = 7.14$, $MSE = .36$, and standard-twice conditions (ARC = .07), $F(1, 72) = 4.04$, $MSE = .36$. Whereas clustering was clearly positive for standard-once subjects, it was at chance for the remaining conditions.[5] Thus, integration appears to create associations between Rp+ and Rp− items, increasing the chance that those items will be interleaved during recall. Such interleaving delays recall of Rp+ items, as seen in the later average position for those items in the integrative-rehearsal (2.35) and standard-twice conditions (2.27), than in the standard-once condition (1.97). These results argue that reduced Rp+ facilitation found for the integrative-rehearsal and standard-twice conditions may reflect increased output interference on those items, arising as a side effect of integration. Rather than posing a problem, reduced facilitation may arise precisely because integration is at work.

The point of this analysis is that when integration alters output dynamics, reduced facilitation may arise, appearing to support an interpretation in terms of differential competition, when it does not. Although we believe that this (and other considerations) renders differential competition unlikely as an account of integration benefits, it would be prudent to address doubts arising from this hypothesis. We pursue those doubts in Experiment 3.

Experiment 3

Experiment 3 addressed two questions. The first concerns the role of Rp+ strengthening in determining integration benefits. In Experiment 2, integrative-rehearsal subjects exhibited less retrieval-induced forgetting than standard-once subjects, but they also showed 18% less Rp+ facilitation, suggesting that those items were less strengthened. Perhaps integration benefits reflect diminished competition by Rp+ items on the final test. Alternatively, reduced facilitation may arise from increased output interference on Rp+ items, brought about by changes in recall order. To test this view, and the role of Rp+ strengthening in producing integration benefits, we controlled test order with a category-stem recall test (e.g., *Fruit O__*). As in prior work (Anderson et al., 1994), subjects were tested on one item at a time, for 7 s each, with tests blocked by category. If variations in Rp+ facilitation in Experiment 2 arose from output interference, controlling test order should equate facilitation across conditions, enabling a clearer assessment of integration effects. This new test format also allowed a manipulation of within-category test order. Rp+ and Rp− items (and also Nrp) could be tested in the first (hereafter, Rp+first, Rp−first, Nrpfirst) or last three positions (hereafter, Rp+second, Rp−second, Nrpsecond items) of their category blocks, providing a measure of output interference.

Our second question concerned the role of study time in reducing retrieval-induced forgetting. Integrative-rehearsal subjects were less impaired than standard-once subjects, but they also got more study time. Perhaps impairment decreases with item strength, regardless of integration. Concern over this alternative is fueled by reduced impairment in the standard-twice condition, which also got extra study time. Although standard-twice subjects reported more spontaneous integration, the effects of integration and study time are hard to disentangle. The solution requires a way of adding study time without increasing spontaneous integration. To achieve this, we included a new control called the focused-rehearsal condition. Subjects in this new condition were given two study passes; however, during the second pass, they silently repeated the item as many times as they could in 7 s, noting repetitions by tick marks. Thus, subjects attended to items for an extra 7 s, but in a way that discouraged integration. If extra study time reduces impairment, less impairment should be found for focused-rehearsal than standard-once subjects. However, if integration, not study time, reduces impairment, both standard-once and focused-rehearsal subjects should show retrieval-induced forgetting, and more than integrative-rehearsal subjects. Less impairment should be found for the latter group, despite similarities to focused-rehearsal subjects in study time and overall recall.

Finally, because increased spontaneous integration reduced impairment in the standard-twice condition of Experiment 2, we should replicate that reduction in Experiment 3, coupled with higher integration rates than in the standard-once or focused-rehearsal conditions. Given higher integration rates, dividing this condition into high and low integrators should reveal less impairment for high integrators.

Method

Subjects

A total of 128 University of Oregon undergraduates participated to fulfill a course requirement. Subjects were randomly assigned to the standard-once, standard-twice, integrative-rehearsal, or focused-rehearsal conditions ($n = 32$ in each case).

[5] This decrease in clustering in both conditions receiving two passes through the study booklet cannot be attributed to a general increase in the likelihood of recalling Rp− items due to increased exposure time. The ARC measure considers the absolute number of items in each category in its computation of above and below chance clustering. As such, the measure is not sensitive to changes in level of recall. To check this, we examined ARC scores for the three conditions holding the relative recall of Rp+ and Rp− items constant. The conclusions from this analysis are exactly the same. For instance, when we focused on categories in which 2 Rp+ and 2 Rp− items were recalled, clustering was .64, −.12, and −.44 respectively for the standard-once, standard-twice, and integrative-rehearsal conditions. The same pattern was evident in all combinations of Rp+ and Rp− scores.

Design

Three factors, study-instructions, retrieval-practice status, and testing order were manipulated. Study instructions was manipulated between-subjects on four levels: standard-once, standard-twice, integrative-rehearsal, and focused-rehearsal. As before, questionnaire ratings were used to make high and low integrator groups within each condition (except for focused-rehearsal subjects, who got a different questionnaire—see *Materials*). Retrieval-practice status was manipulated as in previous experiments.

Because Experiment 3 used a category-stem cued recall final test, we were able to manipulate the test order of exemplars within each category on two levels: the tested-first and tested-second conditions. For each category, items were tested in blocks of six. Tested-first and tested-second items were tested in the first and last three positions of the block, respectively. The dependent measure was the proportion recalled in each condition.

Materials

The materials of Experiment 2 were used, except for the following changes. Test books contained nine categories—one filler, to acquaint subjects with the procedure, followed by eight critical categories. Categories were tested in six-item blocks, one item at a time, by cuing with the category and a letter stem (e.g., *Dog C__*).

To control for output interference, the test position of Rp and Nrp categories was matched. The order of particular categories was also counterbalanced so that the position of every category was equated, across subjects. Finally, the order of items within each category was counterbalanced: For half of the subjects, items a, b, and c were tested in the first category half, and for the remaining subjects, in the second half. This control ensured that all items contributed to test-order and retrieval-practice combinations (e.g., Rp+1st, Rp+2nd, Rp−1st, Rp−2nd, etc.) equally. Crossing our item- and category-position counterbalancing yielded four test orders, which were made into four test books.

The four retrieval-practice counterbalancing conditions were crossed with each test book, yielding 16 practice book–test book combinations. Because 32 subjects were tested in each condition, 2 were tested in every combination within each instructional condition.

Procedure

The procedure of Experiment 2 was used, except for the following changes.

Second study phase. We added instructions to the second study phase to implement the focused-rehearsal condition. Subjects were told to spend the extra 7 s silently repeating the current exemplar as many times as possible. To keep subjects on task, they were asked to make a tick mark under the exemplar for each time they rehearsed it. This task matched their behavior to that of integrative-rehearsal subjects.

Test phase. Subjects were randomly assigned to one of four tests. Subjects were told that each test page contained a category and the first letter of an exemplar (e.g., *Gems D__*). Their task was to recall the study word that fit the stem and to write it underneath the category-stem pair within 7 s, when they would be signaled to proceed.

Results and Discussion

All analyses were done with study instructions, practice counterbalancing, and reported integration (as measured by our questionnaire) as between-subjects factors. Unlike in previous experiments, we found a main effect of Practice Counterbalancing, $F(3, 96) = 2.76$, $MSE = 5.33$, the reasons for which are unclear. However, because this factor did not interact with other variables, it will not be discussed further.

Retrieval Practice

Success rates did not vary reliably between the standard-once ($M = 89\%$), standard-twice ($M = 92\%$), integrative-rehearsal ($M = 93\%$), and focused-rehearsal conditions ($M = 95\%$), $F(3, 96) = 1.75$, $MSE = 11.32$, and also for low ($M = 91\%$) and high integrators ($M = 93\%$), $F(1, 96) = 1.87$, $MSE = 11.32$. The difference between low and high integrators did not interact with Study Instructions, $F < 1$.

Final Recall as a Function of Instructional Condition and Retrieval Practice

In presenting the final recall data, we first discuss our manipulation of Study Instructions; we then consider the effects of testing order and reported integration. Table 4 shows the recall percentages for each item-type in the four conditions. As in Experiment 2, a second study pass

Table 4
Mean Percentage of Words Correctly Recalled as a Function of the Retrieval-Practice Status of an Item and Instruction Condition in Experiment 3

	Retrieval-practice status of an item				
Rehearsal group	Rp+	Rp−	Nrp	Inhibition (Nrp − Rp−)	Facilitation (Rp+ − Nrp)
Standard-once	81	56	63	7	18
Standard-twice	88	70	74	4	14
Integrative-rehearsal	90	72	71	−1	19
Focused-rehearsal	90	63	71	8	19

Note. Rp+ = items given retrieval practice; Rp− = unpracticed items in a practiced category; Nrp = baseline items; standard-once = subjects given one pass through the study list and no special study instructions; standard-twice = subjects given two passes through the study list and no special instructions; integrative and focused-rehearsal = subjects given either integrative- or focused-rehearsal instructions on second study pass.

improved recall, with standard-once subjects recalling less ($M = 66\%$) than subjects in the standard-twice ($M = 76\%$), $F(1, 96) = 9.38$, $MSE = 5.33$, integrative-rehearsal ($M = 76\%$), $F(1, 96) = 8.92$, $MSE = 5.33$, and focused-rehearsal conditions ($M = 74\%$), $F(1, 96) = 5.10$, $MSE = 5.33$. Notably, additional time benefited integrative- and focused-rehearsal subjects equally; thus, variations in impairment across these conditions cannot be attributed to differences in study time or overall recall. These results also replicate the overall facilitation and impairment of Rp+ and Rp− items respectively: Practice improved final recall of Rp+, ($M = 87\%$) relative to Nrp baseline items ($M = 70\%$), $F(1, 96) = 257.57$, $MSE = .74$, but it impaired final recall of Rp− ($M = 65\%$) relative to Nrp items (70%), $F(1, 96) = 13.52$, $MSE = .85$.

Facilitation of practiced items. One of our aims was to equate Rp+ facilitation across conditions by holding within-category recall order constant. As shown in Table 4, facilitation was well matched across the standard-once, standard-twice, integrative-rehearsal, and focused-rehearsal conditions, with the amount not varying by condition (for interactions not involving standard-twice, $F < 1$ in all cases; otherwise, $p > .15$).[6] Comparable Rp+ facilitation effects suggest that the reduced facilitation in Experiment 2 may reflect exaggerated output interference; if integration reduced strengthening, Rp+ facilitation should have been reduced in this study, even with output order controlled.

Variations in retrieval-induced forgetting. Our main questions concerned retrieval-induced forgetting. If integration reduces impairment, more impairment should be found in the standard-once than in the integrative-rehearsal condition. As shown in Table 4, this prediction was confirmed: standard-once subjects were impaired, $F(1, 96) = 8.27$, $MSE = .85$, but integrative-rehearsal subjects were not, $F < 1$, and this difference was reliable, $F(1, 96) = 4.79$, $MSE = .85$. This result replicates the reduced impairment found in Experiments 1 and 2 and extends those findings to a category-stem cued recall test. Crucially, because Rp+ facilitation was well matched across these conditions, less retrieval-induced forgetting for integrative-rehearsal subjects cannot be attributed to reduced competition from Rp+ items.

To see whether adding study time reduced impairment, we compared retrieval-induced forgetting for focused- and integrative-rehearsal subjects, both of whom got increased study time. More impairment was found for focused- than for integrative-rehearsal subjects, $F(1, 96) = 5.50$, $MSE = .85$, despite very similar recall of baseline and practiced items (Nrp and Rp+ = 71 and 90 for both focused and integrative-rehearsal conditions) and also despite similar facilitation (19%). Indeed, focused-rehearsal subjects were impaired $F(1, 96) = 9.59$, $MSE = .85$, and to the same degree as standard-once subjects, $F < 1$. Together, these results show that neither increased study time nor diminished facilitation can explain integration benefits. Integrative-rehearsal subjects show a genuine integration-based reduction in retrieval-induced forgetting.

As shown in Table 4, the standard-twice condition showed a modest, but not reliable, reduction in retrieval-induced forgetting (4%) relative to the standard-once, $F < 1$, and focused-rehearsal conditions, $F(1, 96) = 1.11$, $MSE = .85$. Nevertheless, standard-twice impairment was not reliable, $F(1, 96) = 2.57$, $MSE = .85$, nor did it differ from that found for integrative-rehearsal subjects, $F(1, 96) = 1.67$, $MSE = .85$. Despite this ambiguity, it is likely that this reduced impairment reflects increased spontaneous integration with added study time. We return to this point in our section describing characteristics of questionnaire data.

The Influence of Testing Order on Final Recall Performance

Table 5 shows the recall percentages for the four conditions as a function of retrieval-practice status and testing position within a category block. As expected on the basis of prior work (Anderson et al., 1994), recalling items in the first half of each category block ($M = 75\%$) impaired recall in the second half ($M = 71\%$), $F(1, 96) = 9.59$, $MSE = 1.10$. Interestingly, integrative-rehearsal subjects showed less output interference than the remaining conditions, $F(1, 96) = 4.26$, $MSE = 1.10$. Indeed, output interference was reliable in the remaining conditions, $F(1, 96) = 13.79$, $MSE = 1.10$, (1st − 2nd = 75 − 70 = 5%), but not for integrative-rehearsal subjects, $F < 1$ (78 − 78 = 0%). Thus, it appears that integration reduces output interference in the same way it reduces Rp− impairment. This finding suggests that reduced facilitation in the standard-twice and the integrative-rehearsal conditions in Experiment 2 reflects ceiling effects on Rp+ performance, and not exaggerated output interference, although genuine differences in strengthening cannot be ruled out.

Next we looked at facilitation and retrieval-induced forgetting as a function of testing order. Retrieval-induced forgetting did not interact with test position, $F < 1$, and was reliable for items tested in the first (Nrp1st − Rp−1st = 71 − 67 = 4%), $F(1, 96) = 4.15$, $MSE = .86$, and second halves of a category block (69 − 63 = 6%), $F(1, 96) = 5.88$, $MSE = .86$. It is important to note that this finding replicates prior work showing that retrieval-induced forgetting is not caused by the output of Rp+ items (Anderson et al., 1994). In the standard-once condition, there looks to be more impairment for tested second than tested first items, but this interaction was not reliable, $F < 1$. No other interactions of facilitation, testing order, and instructional condition were reliable.

Characteristics of the Questionnaire Data

As in Experiment 2, standard-once subjects reported less integration ($M = 2.80$) than either standard-twice ($M = 3.31$),

[6] Naturally, these recall data are near ceiling, so one cannot be certain that differences do not exist. However, nothing in the recall data of the integrative-rehearsal and focused-rehearsal conditions even suggests such differences in facilitation. Furthermore, although the integrative-rehearsal condition is closer to ceiling, it shows numerically more facilitation than standard-once, even though the former shows less impairment.

Table 5
Mean Percentage of Words Correctly Recalled as a Function of the Retrieval-Practice Status of an Item, Rehearsal Condition, and Within-Category Testing Position in Experiment 3

	Retrieval-practice status of an item				
Rehearsal group	Rp+	Rp−	Nrp	Inhibition (Nrp − Rp−)	Facilitation (Rp+ − Nrp)
Standard-once					
Tested-first	84	62	65	3	19
Tested-second	78	51	61	10	17
M	81	56	63	7	18
Standard-twice					
Tested-first	92	71	75	4	17
Tested-second	84	69	72	3	12
M	88	70	74	4	14
Integrative-rehearsal					
Tested-first	91	71	71	0	20
Tested-second	89	73	72	−1	17
M	90	72	71	−1	19
Focused-rehearsal					
Tested-first	90	65	74	9	16
Tested-second	90	61	67	6	23
M	90	63	71	8	19

Note. Rp+ = items given retrieval practice; Rp− = unpracticed items in a practiced category; Nrp = baseline items; standard-once = subjects given one pass through the study list and no special study instructions; standard-twice = subjects given two passes through the study list and no special instructions; integrative and focused-rehearsal = subjects given either integrative- or focused-rehearsal instructions on second study pass; tested-first and tested-second = items tested either in the first or second half of a category.

$F(1, 112) = 5.52$, $MSE = .74$, or integrative-rehearsal subjects ($M = 3.61$), $F(1, 112) = 14.10$, $MSE = .74$. Standard-twice and integrative-rehearsal subjects did not differ reliably, $F(1, 112) = 1.97$, $MSE = .74$, again confirming our prediction that with more study time, spontaneous integration would increase. This result fits well with the intermediate impairment found in this condition, as in Experiment 2.

Differences in Recall Performance Across Low and High Integrators

The method for computing integration scores and making low and high groups was similar to that used in Experiment 2, except median splits were done within every combination of learning, practice, and test book for each condition. The rating difference between low ($M = 2.64$) and high integrators (3.52) was clear and similar to that found in Experiment 2 (2.97 vs. 3.73). Differences were clear within the standard-once (2.42 vs. 3.19), standard-twice (2.78 vs. 3.84), and integrative-rehearsal (3.26 vs. 3.96) conditions.[7] Overall recall was similar for low ($M = 74\%$) and high integrators ($M = 72\%$), $F < 1$, and this pattern was similar within study conditions, $F(1, 96) = 1.09$, $MSE = 5.33$.

Surprisingly, in the overall analysis, collapsed over testing order, neither facilitation, $F < 1$, nor retrieval-induced forgetting, $F < 1$, varied across the low and high integrator groups, suggesting that spontaneous integration does not reduce impairment in the category-stem cued recall paradigm. The lack of a difference in overall impairment, however, masks an interaction between impairment, reported integration, and testing order, $F(1, 96) = 6.59$, $MSE = 1.45$. For items tested first in their category blocks, the predicted reduction in retrieval-induced forgetting was observed, with less retrieval-induced forgetting for high (Nrp − Rp− = 67 − 70 = 3% facilitation) than for low integrators (73 − 66 = 7%), $F(1, 96) = 5.91$, $MSE = 1.09$. However, when items were tested second, impairment did not differ reliably, $F(1, 96) = 1.72$, $MSE = 1.21$. This result suggests a disparity between spontaneous and instructed integration, with the latter showing impairment in both test positions.

To look at the generality of this finding, we analyzed integration reports in six prior studies that used a single study pass and a category-stem recall test. All used the questionnaire employed in this experiment, which was used to make groups of 91 low and 90 high integrators. Figure 1 presents the results of this analysis for the two materials sets used. As expected, little impairment was found for high integrators (Nrp − Rp− = 61 − 61 = 0%), $F < 1$, but low integrators were impaired, (61 − 54 = 7%), $F(1, 133) = 23.86$, $MSE = 1.04$, and this interaction was reliable, $F(1, 133) = 12.47$, $MSE = 1.04$. High integrators were unimpaired, despite facilitation (24%), similar to that seen for low integrators (22%), $F(1, 133) = 1.29$, $MSE = .95$. Notably, benefits were similar across test positions: Low integrators

[7] Because focused-rehearsal subjects were not asked about integration, we focused on the data from the remaining conditions.

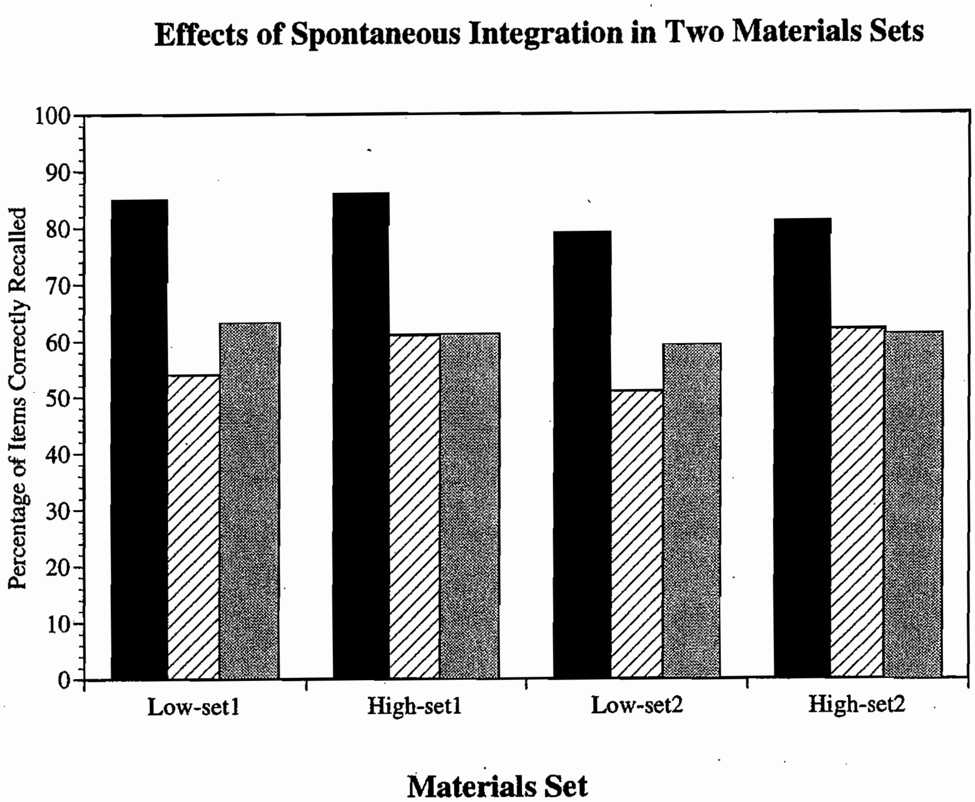

Figure 1. A meta-analysis of the effects of spontaneous integration on retrieval-induced forgetting in six previous studies ($N = 181$). All studies used a category-stem cued recall final testing procedure. The left side ($n = 51$ and 50 for low and high integrators, respectively) and the right side ($n = 40$ for both low and high integrators) of the figure reflect two different materials sets. *Low* and *high integrators* refer to subjects reporting a low versus a high degree of integration on our postexperimental questionnaire. The data reported are the mean number of words correctly recalled for Rp+ (practiced items [black bars]), Rp− (unpracticed members of a practiced category [striped bars]), and Nrp (baseline [gray bars]) items.

were impaired for items tested first (Nrp − Rp− = 64 − 55 = 9%), $F(1, 133) = 14.31$, $MSE = 1.30$, and second in their categories (58 − 52 = 6%), $F(1, 133) = 5.61$, $MSE = 1.30$, and high integrators showed little impairment in either the first (63 − 60 = 3%), $F(1, 133) = 1.08$, $MSE = 1.30$, or second halves (58 − 62 = 4% facilitation), $F(1, 133) = 1.37$, $MSE = 1.35$. Thus, integration reports in these studies yielded results mirroring the pattern found in the manipulation of Experiment 3. These findings argue that variations in the benefits of spontaneous integration across test positions in Experiment 3 may be unrepresentative. They also provide further evidence that integration benefits are not caused by differential competition from Rp+ items.

Summary

Experiment 3 replicates the major results of Experiments 1 and 2 and provides specific evidence that integration protects against retrieval-induced forgetting. Integrative-rehearsal subjects were significantly less impaired than focused-rehearsal subjects, despite the matching of these groups for study time, overall recall, Rp+ facilitation, and the presence of rehearsal. The main difference between these conditions was in the integrative nature of rehearsal, which completely eliminated impairment. Thus, the findings of Experiment 3 argue strongly that integration benefits are not caused by differential study time or differential competition but rather by a distinct protective effect of integration.

General Discussion

Previous work has shown that recalling information from long-term memory impairs the long-term retention of semantically related representations (Anderson et al., 1994). The present experiments show that whether or not a related representation will be impaired depends critically on the degree to which it has been integrated with the retrieval target. If the representations associated to a retrieval cue have few interconnections with each other, retrieving one of those representations will impair retention of the remaining ones. However, if the representations are interlinked or integrated with one another, as well as to the cues guiding retrieval, recalling any one associate can proceed with little cost to the remainder. This conclusion is supported by three main findings.

First, in all three experiments, we found that instructing subjects to rehearse the members of a category during an initial study phase reduced retrieval-induced forgetting. We found reduced impairment on a category-cued recall test in Experiments 1 and 2 and on a category-stem cued recall test in Experiment 3. Although spending time on rehearsal may also reduce subjects' attention to category-exemplar associations (and increase the amount devoted to exemplars themselves), Experiments 2 and 3 show that neither of these factors reduced retrieval-induced forgetting. Rather, the current effects were specifically caused by the integrative quality of rehearsal; reductions in forgetting were largest when rehearsal instructions directed people to interrelate exemplars and were smallest when they focused people on individual items. These findings argue that integrating the associates of a retrieval cue renders those items less vulnerable to retrieval-induced forgetting, despite the role of suppression in that form of impairment (Anderson et al., 1994; Anderson & Spellman, 1995).

Second, all three experiments found that even when people were not explicitly asked to integrate items, they frequently did so on their own. All experiments used a final questionnaire to assess how often people integrated during the study phase. When we categorized subjects into high and low "spontaneous integrators," we found that self-initiated integration had the same effect on retrieval-induced forgetting as we found with our instructions: less retrieval-induced forgetting on both the category cued and category-stem cued recall tests. Because all of the questionnaires used in the present studies specifically asked about integration, the reduced impairment shown by high integrators supports the idea that integration protects people from retrieval-induced forgetting. These data also echo those found in previous work by Martin and Dean (1964), who found that subjects' introspective integration reports reliably predicted proactive interference.

Third, in both Experiments 2 and 3 there was a trend toward less retrieval-induced forgetting when initial study time was increased. However, these reductions do not appear to arise from increased study time per se. Rather, adding study time reduces impairment by increasing the time for spontaneous integration. In both Experiments 2 and 3, people receiving extra study time gave higher integration ratings on the final questionnaire, even when they were not asked to integrate. These increases in ratings were substantial, approaching the levels reported by subjects who were explicitly instructed to rehearse. Less retrieval-induced forgetting would be expected if extra study time increased spontaneous integration. This explanation of the effects of study time is reinforced by two findings. First, low spontaneous integrators in the standard-twice condition of Experiment 2 showed just as much retrieval-induced forgetting as did standard-once subjects, even though the latter subjects received less study time. If adding study time reduced impairment, standard-twice subjects should have shown less impairment, regardless of their reported rate of integration. Second, subjects who were given extra study time in Experiment 3, but who were discouraged from integration by focused rehearsal instructions, failed to show reduced impairment. Thus, adding study time does not by itself reduce impairment; rather, when people receive more study time in this intentional learning procedure, they are more likely to integrate and thus show less retrieval-induced forgetting.

One feature of the present findings we did not expect was the reduced facilitation of practiced items found in Experiment 2. In Experiment 2, subjects who were asked to integrate showed 19% less facilitation than standard-once subjects. Standard-twice subjects also showed reduced facilitation. These cases of reduced facilitation occurred despite high retrieval-practice success rates in all conditions, showing that differences in practice success are not responsible. These groups did differ on another dimension, however; when people integrated items, they were far more likely to interleave their final recall of practiced and unpracticed exemplars, presumably because recalling practiced items triggered recall of interrelated unpracticed items. Interleaving recall in this way increased the output position for practiced items, suggesting that those items may have suffered exaggerated output interference. This view was not supported by Experiment 3, which established that integration reduces output interference in the same way as it reduces Rp− impairment. It thus seems unlikely that exaggerated output interference could have reduced Rp+ facilitation in Experiment 2. Whatever caused this reduced facilitation, Experiment 3 shows that reduced facilitation is not responsible for the integration benefits observed for retrieval-induced forgetting.

The present findings extend the wealth of prior research on integration effects by underscoring the importance of integration as a boundary condition on retrieval-induced forgetting. The study of integration in this new empirical context is important for both theoretical and practical reasons. Because retrieval-induced forgetting is believed to be caused by active suppression processes that arise during recall (Anderson et al., 1994; Anderson & Spellman, 1995; Blaxton & Neely, 1983; Dagenbach, Carr, & Barnhardt, 1990), it is unclear whether the integration benefits typically enjoyed by traditional associative interference processes would affect this phenomenon in the same way. Indeed, although integration might reduce suppression in several ways (a point to which we return in the next section), there are some analyses that would lead one to expect increased impairment. For example, unlike fan interference, retrieval-induced forgetting is caused by the prior recall of target items from incomplete retrieval cues. It is possible that the indirect connections between competitors resulting from integration might heighten the activation of those items during retrieval practice in a way that is particularly disruptive to recall. Such disruption might increase the demand for suppression, particularly to the extent that integration makes competing items less discriminable (Dagenbach et al., 1990). Because such alternative views are possible, it is important to evaluate the impact of integration on retrieval-induced forgetting directly. The present findings, which show that this latter theoretical analysis does not extend to the current situation, thus add an important empirical constraint that must be met by any theoretical treatment of recall-based inhibition.

On a practical level, the current findings generalize the range of situations in which integration benefits may be expected to occur. Despite the abundance of studies on integration benefits, no work has examined whether such effects would be affected by uneven experience with material after it has been integrated. Yet, much of the integrated knowledge we acquire is used unevenly; in everyday settings, people frequently access only that subset of their knowledge about a topic that is consistently useful. Although integration clearly helps when all of the information is to be maintained, there is ample precedent from research on part-set cuing and retrieval-induced forgetting to suggest that uneven use of the information may eliminate or even reverse the benefits of organization. Nevertheless, the present findings show that integration benefits persist in the face of such uneven use and occur despite the role of inhibitory processes in active recall. These findings thus greatly strengthen the case for integration as a general factor that alleviates forgetting from long-term memory.

When considering how integration might affect retrieval-induced forgetting, we reviewed three bodies of research: research on fact retrieval, mediated transfer, and part-set cuing. Each of these bodies of work suggests mechanisms that could have influenced our results. In the next three sections, we argue that two mechanisms suggested by this work—diminished suppression and mediated retrieval—contributed to our findings, but that strategy disruption effects, sometimes seen in research on part-set cuing, did not. We then discuss how understanding the role of integration as a boundary condition on retrieval-induced forgetting might constrain how retrieval processes shape the accessibility of knowledge and experience. In particular, we discuss limits on forgetting in three situations: retrieval from complex knowledge representations, the retrieval characteristics of experts, and retrieval from within individual episodic memories.

Diminished Suppression

One approach to explaining reductions in retrieval-induced forgetting builds on the finding that integration can reduce retrieval competition. One can learn many new facts about a concept with little cost to retrieval speed, if those facts are integrated (Moeser, 1977, 1979a, 1979b; Myers et al., 1984; Radvansky, Spieler, & Zacks, 1993; Radvansky, Wyer, Curiel, & Lutz, 1997; Radvansky & Zacks, 1991; Smith et al., 1978). If integrating the members of a category also reduces competition, there should be little need to suppress nontargets during retrieval practice. If retrieval-induced forgetting reflects the enduring effects of suppression, as indicated by prior work (Anderson et al., 1994; Anderson & Spellman, 1995), integration should insulate nontargets from retrieval-induced forgetting. Thus, integration benefits might reflect the dependence of inhibition on competitor interference (Anderson et al., 1994; Anderson, E. Bjork, & R. Bjork, 1993; Anderson & Spellman, 1995).

The diminished suppression hypothesis accounts for the major findings of the present study. It explains why both spontaneous and instructed integration reduce retrieval-induced forgetting and focused rehearsal does not. But perhaps its strongest feature is that it explains why integration benefits do not depend on within-category testing position. In the final tests of Experiments 1 and 2, subjects could recall items in any order they wished. In Experiment 3, however, we manipulated testing order with a category-stem cued recall test. If integration benefits were produced solely by mediated retrieval (see next section), one might have expected greater benefits for unpracticed items when practiced items were tested first. It is important to note that unpracticed items were less impaired in both test positions, showing that integration benefits do not depend on the prior recall of practiced items. This finding is easily explained if integration eliminated the need to suppress competitors during retrieval-practice. Further work using the independent cue method (Anderson & Spellman, 1995) must be conducted, however, before we can determine whether these effects reflect reduced suppression.

Mediated Retrieval

A second approach to explaining integration benefits builds on mechanisms thought to be at work in classical studies of verbal mediation. Research on verbal mediation assumed that people could use connections between competing associates as indirect retrieval routes by which they could circumvent interference (for reviews see Horton & Kjeldergaard, 1961; Jenkins, 1963; Kjeldergaard, 1968; Postman, 1971). In the context of the present materials, if cat and dog were interrelated while being studied as members of animal, later recall of one exemplar should cue retrieval of the other. Although such connections should benefit recall in both practiced and baseline categories, inhibited items may benefit more because the strengthening of their practiced category-mates should greatly improve access to interexemplar pathways. For instance, even if dog and cat compete when animal is presented as a cue, and even if cat is suppressed, practicing dog may compensate by enhancing access to the dog–cat link. Thus, integration could simply offset or mask the suppression that normally causes retrieval-induced forgetting.

Two observations suggest that at least some of the present integration benefits stems from the use of indirect retrieval pathways. First, given that people were either instructed to integrate, or that they did so as a deliberate strategy, it seems likely that they would have taken advantage of their efforts on the final test. If people used their integration strategies during recall, reduced impairment would arise from mediated retrieval. However, because we did not include questionnaire measures to address subjects' test strategies, this argument must remain plausible speculation. Second, in the final recall phase of Experiment 2, subjects interleaved practiced and unpracticed exemplars to a greater extent in the integrative-rehearsal and standard-twice conditions than in the standard-once condition. This occurred even though the items that were interleaved between practiced items were very likely to be less strongly associated to the category cue than the postponed practiced items. This recall pattern suggests that subjects may have used practiced items as cues to recall associated unpracticed items. If reduced impairment was caused solely by diminished suppression, and not by mediated retrieval, it is unclear why such interleaving would have occurred. Nevertheless, the present findings are compatible with both approaches, and it is possible that integration benefits reflect some combination of reduced suppression and mediated retrieval.

Strategy Disruption

This study found no evidence suggesting that retrieval practice causes strategy disruption (Basden & Basden, 1995). Basden and Basden (1995) found that presenting several category exemplars as cues disrupted recall of remaining items when those cues conflicted with the organization that subjects had developed over several study-test cycles. Their finding shows that when subjects integrate exemplars, their recall performance may be particularly sensitive to disruption by exposure to a subset of the studied items. Although the present studies were not designed to test the strategy disruption hypothesis, one can imagine how such dynamics might have influenced our results: Instructing people to integrate encourages organization of the study items, and retrieval practice involves repeated reexposure to a subset of those organized items. To the extent that the subset of practiced items was incongruent with the subjects' organization, one should have seen evidence of strategy disruption, particularly in the integrative-rehearsal condition. That integration reduced retrieval-induced forgetting thus contradicts this extension of Basden and Basden's analysis to retrieval-induced forgetting.

There are a number of reasons why this study may have failed to yield evidence for strategy disruption. One possibility is that the method of integration used by our subjects did not yield a true retrieval strategy. Our instructions focused people on forming simple interrelationships between individual items. Perhaps the strategies developed in the repeated study-test procedure of Basden and Basden (1995) were more complex and vulnerable to disruption. Although this possibility is reasonable, the strategy disruption notion was meant as a general account of part-set cuing impairment. Because part-set cuing studies typically use a single brief study exposure, strategy formation cannot depend on a repeated study-test trials procedure if it is to explain that phenomenon. If a brief exposure is enough to yield disruptable retrieval strategies, our study procedure should have been sufficient.

A more likely explanation of our failure to see greater impairment for high integrators stems from differences in how subsets are reexposed in the two procedures. Part-set cuing subjects usually get the entire subset of cues presented to them at test, encouraging the reorganization proposed to cause strategy disruption. In contrast, when subjects do retrieval practice, they focus on a single exemplar at a time, with tests of other items in that category separated by test trials for entirely different categories. Thus, retrieval-practice may make it less likely for reorganization to occur. Nevertheless, evidence of strategy disruption in other settings suggests that integration, though useful in reducing competition and retrieval-induced forgetting, may create problems of its own.

Implications of Integration for the Stability of Knowledge

Regardless of the mechanisms that underlie integration benefits, the current findings argue that this factor strongly constrains when retrieval-based inhibitory processes will cause forgetting. One interesting implication of this constraint is that complex knowledge structures that might otherwise be vulnerable to inhibition may be protected by the stability conferred by integration. Because we depend on consistent access to such structures, integration may prevent some rather paradoxical effects of inhibition. In this section, we describe two situations in which integration may be crucial in serving this function: Retrieval from conceptual representations and situation models, and the retrieval of domain knowledge by experts. We then describe how both retrieval-induced forgetting and integration may shape what we eventually remember about individual episodes.

Retrieval-Induced Forgetting and Complex Knowledge Structures

Complex knowledge representations such as concepts, schemas, and situation models, are made from simpler parts that contribute to their meaning. According to the simple view of retrieval-induced forgetting, selectively recalling particular aspects of these representations should suppress the remaining parts, eroding access to much of the structure. Such erosion seems at odds with both intuition and the very purpose of constructing such structures: To ensure that knowledge relevant to an entity or event becomes available to the system. If access to the parts of representations is to be preserved in the face of suppression, there must be one or more processes that counteract or prevent this erosion.

One factor likely to diminish these effects is the tendency for the elements of complex knowledge structures to recur together whenever the referent of the representation is encountered in the world. Such reexposure to constituent features should preserve the representation. However, there are circumstances in which reexposure to a representation's parts may not be guaranteed. During discourse or reading, for instance, concept terms occur in the absence of their referents, making it possible that, without reexposure, selective emphasis of some features may entail suppression of others. Furthermore, the features of abstract concepts may never be reencountered after they are learned, perhaps making them particularly subject to retrieval-based suppression. How do we retain access to concepts in such cases? Perhaps integration adds stability. If the present findings are representative, connections among the components of a representation should allow some elements to be recalled without suppressing the remainder.

Two lines of work suggest that integration confers the stability necessary to protect complex representations from suppression. The first concerns how properties are retrieved from conceptual representations. According to traditional views of interference, the more features involved in a concept, the more competition there should be while retrieving a feature. This perspective argues that complex concepts may be quite unstable. However, this prediction does not consider connections between features, and the indirect source of facilitation they provide. Kroll and Klimesch (1992) examined the effect of feature-feature integration, using category exemplars (e.g., lemon, banana, orange) varying in the preexperimental number, strength, and connectivity of features. In a semantic verification task, they found that subjects were faster to verify properties (e.g., apple-red) of concepts that had more features than of concepts that had less, provided those features were interrelated. Thus, as the authors predicted, competition can be reversed when features are integrated. Interestingly, Kroll and Klimesch (1992) also found that as the number of features in a concept increased, so too did the average rating of associative strength between the features. This finding suggests the intriguing possibility that as concepts grow in complexity, integration may take on a prominent role in preserving stability. If the integration effects in their study are representative, recalling features from natural concepts may often proceed without competition, and thus without retrieval-induced forgetting.

The second line of work relevant to the stabilizing effects of integration concerns the retrieval of information from situation models. A large body of work on discourse comprehension has shown that people construct mental models of situations described during a discourse (see Zwaan & Radvansky, 1998, for a review). Such models provide a framework for interpreting how the facts being presented interrelate with one another. If discourse comprehension relies on situation models, it is important to understand the mechanisms of retrieval from such representations. According to work on fact retrieval (Anderson, 1974, 1983), one might expect that adding component facts to such a model should increase retrieval competition. If inhibitory processes resolve competition, as suggested by Anderson and Spellman (1995), then as situation models grow, they should become highly vulnerable to retrieval-induced forgetting, and thus quite unstable. Simply retrieving one aspect of the model during comprehension should suppress the remainder, functionally dismantling the structure needed to guide discourse processing.

Studies by Radvansky and colleagues have shown, however, that when facts about an object conform to a single situation model, they no longer compete (Radvansky & Zacks, 1991; Radvansky et al., 1993, 1997). Many new facts can be learned about a situation with little cost to retrieval time. Radvansky has argued that when facts concern a single situation, they get integrated into a model that functions as a unit. Because the model integrates component facts, no competition should arise. The view presented here is compatible with that one, although we would emphasize two points. First, integration benefits do not require the construction of situation models per se, but occur whenever the items associated to a cue are linked. Second, we argue that integration reduces both competition and the suppression that would arise from resolving competition. Consistent with that idea, Radvansky has shown that recognizing a fact impairs subsequent recognition of related facts, but only when those facts are not part of an integrated model (Radvansky, in press). These findings argue that integration protects retrieval within a situation model from both competition and retrieval-based suppression, affording stability vital to the role of such representations in sustaining discourse comprehension.

Retrieval-Induced Forgetting and the Paradox of Expertise

The dilemma that arises when we consider how retrieval-induced forgetting might affect complex knowledge representations also arises when we consider retrieval of general knowledge about a domain. If learning more about a concept always increased competition, we would face the odd conclusion that, as we acquire more expertise in a domain, we should be less adept at retrieving our knowledge. Furthermore, if inhibitory processes resolve competition, experts should be very prone to retrieval-induced forgetting. Yet, experts seem to have better, not worse, access to knowledge in their domains. Thus, retrieval-induced forgetting, like interference generally, confronts what Smith et al. (1978) called the "paradox of expertise." How can experts know more about their domains and not suffer heightened interference and—more important for our purposes—greater retrieval-induced forgetting?

As a solution to this paradox, Smith et al. (1978) proposed that the superior memory of experts for knowledge in their domains may arise from integration. According to this view, experts do not simply learn new facts as isolated associations to a topic, but rather interpret them in terms of a rich body of background knowledge. Interpreting a fact in this way should integrate it with other facts, eliminating interference during retrieval. To illustrate this, Smith et al. showed that several newly studied facts that could be unified by a theme or script did not exhibit fan interference. Although these benefits may sometimes be produced by changes in subjects' retrieval strategy for thematically related materials (Reder & Anderson, 1980; Reder & Ross, 1983), the finding that integration reduces competition was borne out in many later studies in which those strategies are unlikely (Myers et al., 1984; Radvansky & Zacks, 1991; Radvansky et al., 1993). Because a deep understanding of a domain is likely to entail knowledge of the interrelationships between many facts, it is plausible that integration protects experts from retrieval competition. Thus, Smith et al.'s solution to the paradox of expertise may be correct, at least for retrieval competition.

Fortunately, if Smith et al.'s (1978) hypothesis about competition is correct, experts should also show less retrieval-induced forgetting. If integration reduces competition, experts should be able to recall facts in their domain without disrupting related knowledge. This benefit may arise either because integration reduces the need to suppress related knowledge, or because their intricate webs of interrelated facts provide many ways in which suppressed knowledge may be reactivated. It should be noted, however, that even experts may suffer retrieval-induced forgetting. Although a high degree of integration seems likely for the expert, there are probably limits (in the content material itself) to what can be integrated, particularly between large subareas of a domain that have no obvious relations. During periods of prolonged focus, in which experts become specialists in some issue, even they may lose access to material with which they were once adept.

Retrieval-Induced Forgetting and Autobiographical Memories

A final limitation suggested by our findings concerns the occurrence of retrieval-induced forgetting within individual episodic memories. Most studies of retrieval-induced forgetting have looked at the effects of retrieving one episode on the recall of another, rather than the effects of retrieval within a single episode. For instance, most studies focus on how retrieving some exemplars alters the ability to recall others in the category. Although all of these exemplars are studied within a single session, it is assumed that items have distinct episodic representations reflecting their presentation on different study trials. Might the mechanisms that suppress competing episodes also operate within an episode? For instance, will recalling some aspects of a particular autobiographical memory inhibit other aspects?

The present analysis suggests that for many memories, the answer may be no. To the extent that events instantiate well-learned schemas, they may have internal structure that links component events together. This structure should reduce competition, and thus the need for inhibition. Even if an episode is not schematic, component actions may be linked by the causal and explanatory relations that people infer while understanding events (Myers et al., 1984). If components are integrated in this way, little retrieval-induced forgetting should occur. However, there are conditions under which within-event inhibition might still occur (Conway, 1996, 1997). Those parts of an episode that are not well integrated should remain vulnerable to retrieval-induced forgetting. Unless these isolated associates get integrated, they may be suppressed as other parts of the episode are recalled. Suppressing the parts that "don't fit," should gradually sculpt the memory so that it conforms to the structure of the preexisting schema or schemas of which it is an instance. Thus, the schematization of memories over time, noted by Bartlett (1932) may be achieved, in part, by the retrieval-induced forgetting that arises from selectively retrieving the schema-consistent aspects of experience.

Concluding Remarks

Prior work on retrieval-induced forgetting has focused on establishing the disruptive effects of recalling information from long-term memory. Although the observation that remembering can cause forgetting is important, we must also understand the limits of this process. There are many conditions in which we do not want the alternative representations associated to a retrieval cue to be inhibited while we recall a target item. For instance, when the alternatives are parts of a larger structure, it seems in our best interests to preserve all of those parts if we wish to retain the ability to represent objects and events in the world. If retrieving an item always involves inhibitory processes, how do we retain consistent access to all of the parts of complex representations?

One likely answer to this question lies in an old idea about how to prevent forgetting: integration. Whenever the items associated to a retrieval cue are linked with one another, their tendency to compete is reduced. This reduced competition has been found in classical studies of proactive and retroactive interference, as well as in modern research on the retrieval of factual and conceptual knowledge. A variety of interrelationships produce integration benefits, including those based on simple associations, propositions, or even complex structures such as scripts or situation models. These findings argue that integration may create a mutually reinforcing representation in which associates "cooperate" rather than compete. The current results suggest the additional point that this cooperation may yield another benefit—a benefit arising from the reduced need to inhibit competing items. Regardless of how integrating representations reduces impairment, the present findings illustrate the importance of integration in constraining how retrieval processes shape what we ultimately retain from past experience.

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, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology, 6,* 451–474.

Anderson, J R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (1993, November). *Strengthening is not enough: Evidence against the blocking theory of retrieval inhibition.* Paper presented at the 34th Annual Meeting of the Psychonomic Society, Washington, DC.

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., & 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.

Bartlett, F. C. (1932). *Remembering.* Cambridge, England: Cambridge University Press.

Basden, D. R., & Basden, B. H. (1995). Some tests of the strategy disruption interpretation of part-list cuing inhibition. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition. 21, 1656–1669.

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut norms [Monograph]. *Journal of Experimental Psychology, 80,* 1–46.

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.

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

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology, 49,* 229–240.

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

Ciranni, M. A., & Shimamura, A. P. (1998). *Retrieval-induced forgetting in episodic memory.* Manuscript submitted for publication.

Conway, M. A. (1996). Autobiographical memories and autobiographical knowledge. In D. C. Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 67–93). Cambridge, England: Cambridge University Press.

Conway, M. A. (1997). Past and present: Recovered memories and false memories. In M. A. Conway (Ed.), *False and recovered memories* (pp. 150–191). Oxford, England: Oxford University Press.

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.

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

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.

Horton, D. L., & Kjeldergaard, P. M. (1961). An experimental analysis of associative factors in mediated generalization. *Psychological Monographs: General and Applied, 75,* 1–26.

Horton, D. L., & Wiley, R. E. (1967). The effect of mediation on the retention and strength of previously formed associations. *Journal of Verbal Learning and Verbal Behavior, 6,* 36–41.

Jenkins, J. (1963). Mediated associations: Paradigms and situations. In C. N. Cofer & B. S. Musgrave (Eds.), *Verbal behavior and learning: Problems and processes* (pp. 210–257). New York: McGraw-Hill.

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

Kjeldergaard, P. M. (1968). Transfer and mediation in verbal learning. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory* (pp. 1071–1075). Englewood Cliffs, NJ: Prentice Hall.

Kroll, N. E., & Klimesch, W. (1992). Semantic memory: Complexity or connectivity? *Memory & Cognition, 20,* 192–210.

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

Martin, R. B., & Dean, S. J. (1964). Implicit and explicit mediation in paired-associate learning. *Journal of Experimental Psychology, 68,* 21–27.

McCloskey, M., & Bigler, K. (1980). Focused memory search in fact retrieval. *Memory & Cognition, 8,* 253–264.

McEvoy, C. L., & Nelson, D. L. (1982). Category name and instance norms for 106 categories of various sizes. *American Journal of Psychology, 95,* 581–634.

McGeoch, J. A. (1936). Studies in retroactive inhibition: VII. Retroactive inhibition as a function of the length and frequency of presentation of the interpolated lists. *Journal of Experimental Psychology, 19,* 674–693.

Melton, A. W., & Irwin, J. M. (1940). The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *The American Journal of Psychology, 3,* 173–203.

Moeser, S. D. (1977). Recognition processes in episodic memory. *Canadian Journal of Psychology, 31,* 41–70.

Moeser, S. D. (1979a). Retrieval interference in the independent storage of related episodic traces. *Canadian Journal of Psychology, 33,* 182–192.

Moeser, S. D. (1979b). The role of experimental design in investigations of the fan effect. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 125–134.

Mueller, C. W., & Watkins, M. J. (1977). Inhibition from part-set cueing: A cue-overload interpretation. *Journal of Verbal Learning and Verbal Behavior, 16,* 699–710.

Myers, J. L., O'Brien, E. J., Balota, D. A., & Toyofuku, M. L. (1984). Memory search without interference: The role of integration. *Cognitive Psychology, 16,* 217–242.

Newell, A., & Rosenbloom, P. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition.* Hillsdale, NJ: Erlbaum.

Nickerson, R. S. (1984). Retrieval inhibition from part-set cuing: A persisting enigma in memory research. *Memory & Cognition, 12,* 531–552.

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

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.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88,* 93–134.

Radvansky, G. A. (in press). Memory retrieval and suppression: The inhibition of situation models. *Journal of Experimental Psychology: General.*

Radvansky, G. A., Spieler, D. H., & Zacks, R. T. (1993). Mental model organization. *Journal of Experimental Psychology, 19,* 95–114.

Radvansky, G. A., Wyer, R. S., Curiel, J. M., & Lutz, M. F. (1997). Situation models and abstract ownership relations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1233–1246.

Radvansky, G. A., & Zacks, R. T. (1991). Mental models and the fan effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 940–953.

Reder, L. M., & Anderson, J. R. (1980). A partial resolution of the paradox of interference: The role of integrating knowledge. *Cognitive Psychology, 12,* 447–472.

Reder, L. M., & Ross, B. H. (1983). Integrated knowledge in different tasks: The role of retrieval strategy on fan effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 55–72.

Roediger, H. L. (1973). Inhibition in recall from cueing with recall targets. *Journal of Verbal Learning and Verbal Behavior, 12,* 644–657.

Roediger, H. L. (1978). Recall as a self-limiting process. *Memory & Cognition, 6,* 54–63.

Roenker, D. L., Thompson, C. P., & Brown, S. C. (1971). Comparison of measures for the estimation of clustering in free-recall. *Psychological Bulletin, 76,* 45–48.

Russell, W., & Storms, L. (1955). Implicit verbal chaining in paired-associate learning. *Journal of Experimental Psychology, 49,* 287–293.

Rundus, D. (1973). Negative effects of using list items as recall cues. *Journal of Verbal Learning and Verbal Behavior, 12,* 43–50.

Schooler, J. W., Fiore, S. M., & Brandimonte, M. A. (1997). At a loss for words: Verbal overshadowing of perceptual memories. *The Psychology of Learning and Motivation, 37,* 291–339.

Schulz, R., Liston, J., & Weaver, G. (1968). The a-b, b-c, a-c mediation paradigm: Recall of a-b following a-c learning. *Journal of Verbal Learning and Verbal Behavior, 7,* 602–607.

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

Slamecka, N. J. (1968). An examination of trace storage in free recall. *Journal of Experimental Psychology, 76,* 504–513.

Smith, E. E., Adams, N., & Schorr, D. (1978). Fact retrieval and the paradox of interference. *Cognitive Psychology, 10,* 438–464.

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

Tulving, E. (1962). Subjective organization in free recall of unrelated words. *Psychological Review, 69,* 344–354.

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

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123,* 162–185.

Appendix

Categories and Exemplars Used in Experiments 1, 2, and 3, Divided Into the Four Practice Counterbalancing Sets (AS1, AS2, BS1, BS2)

Category	Exemplar Set S1	Exemplar Set S2
Set A		
Gems	Ruby, opal, topaz	Diamond, emerald, jade
Instruments	Banjo, guitar, saxophone	Oboe, clarinet, flute
Dogs	Beagle, setter, Labrador	Husky, collie, terrier
Cloth	Velvet, burlap, linen	Wool, polyester, silk
Set B		
Tools	Drill, screwdriver, chisel	Wrench, hammer, pliers
Flowers	Violet, tulip, geranium	Daisy, orchid, carnation
Vegetables	Broccoli, cucumber, lettuce	Pea, radish, asparagus
Wood	Cedar, birch, maple	Walnut, spruce, pine

Received March 3, 1998
Revision received September 23, 1998
Accepted September 23, 1998 ■