Journal of Experimental Psychology: General
2007, Vol. 136, No. 2, 200–216

0096-3445/07/$12.00 DOI: 10.1037/0096-3445.136.2.200

Probabilistic Retroactive Interference: The Role of Accessibility Bias in Interference Effects

Larry L. Jacoby, Anthony J. Bishara, Sandra Hessels, and Andrea Hughes
Washington University in St. Louis

Probabilistic retroactive interference (RI) refers to the interfering effects of intermixing presentations of an earlier studied response (A-B) with presentations of a competing response (A-D). As an example, for a 2/3 condition, a cue word was presented with its earlier studied response twice and its competing response once during the interference phase. Performance on direct and indirect tests of memory for earlier studied responses was combined to reveal dissociations between effects on recollection and accessibility bias. Manipulating probabilistic RI influenced accessibility bias but left recollection unchanged. Effects of probabilistic RI were compared with effects of traditional, nonprobabilistic RI. The authors contrast their dual-process model with traditional accounts of RI and discuss the importance of distinguishing between recollection and accessibility bias for understanding interference effects.

Keywords: retroactive interference, accessibility bias, recollection, process dissociation

"What grade did you obtain in your first high school math course?" Answers to this question are likely to be distorted by memory for grades obtained in subsequent math courses (Bahrick, Hall, & Berger, 1996). For example, if unable to remember the grade for the particular course, people who usually got an A in subsequent math courses might incorrectly guess that they obtained an A in their first course even if their actual grade was lower. Effects of this sort can be described as being produced by probabilistic retroactive interference (RI) because of the nonzero probability of the grade obtained in the first course being among those earned in subsequent courses. The probabilistic nature of the interference contrasts with the traditional procedure of investigating RI by requiring participants to first learn a list of paired associates (List 1), followed by a second list (List 2), with the cues remaining constant across lists but the responses changing (A-B, A-D). To produce probabilistic RI, presentations of A-B would be intermixed with presentations of A-D in List 2. Such probabilistic change is common, as in the example of high school grades, and likely more common than is discrete change of the sort used in traditional investigations of RI.

There are numerous examples of probabilistic interference effects in the literature although they are not described as being such. Ross (1989) reviewed results showing that people rely on implicit theories to construct their personal history and that reliance on theories can exaggerate the consistency between the present and past. For example, McFarland and Ross (1987) assessed university students' evaluation of their dating partners and found that participants underestimated the amount of change that occurred over a few months. In the language of probabilistic RI, memory for an earlier held attitude was interfered with by fluctuations in the attitude during the intervening months. In this vein, recall errors resulting from reliance on a schema (e.g., Bartlett, 1932; Brewer, 2000) can also be thought of as originating from probabilistic interference in that a schema includes aspects of an event that usually occur but do not always do so. As typically investigated, schema effects constitute an example of probabilistic proactive interference (PI)—an effect of what has usually happened in the past on memory for a later, particular event—but schemas can also originate from subsequent events and, so, serve as a source of probabilistic RI.

We argue that a dual-process model that distinguishes between recollection and accessibility bias provides a useful perspective for understanding probabilistic RI as well as PI (e.g., Jacoby, Debner, & Hay, 2001). For recollection, retrieval is cognitively controlled and tightly constrained by effortful reinstatement of study context, whereas accessibility bias relies on a more automatic, less constrained use of memory. We manipulated probabilistic RI and examined effects on recollection and accessibility bias. As is described below, our recollection/accessibility bias distinction is analogous to a distinction between discriminability and response bias (e.g., Snodgrass & Corwin, 1988).

Although it is common to separate effects on response bias from effects on discriminability for tests of recognition memory (e.g., Snodgrass & Corwin, 1988), such a distinction is not made by classic interference theory (e.g., Postman & Underwood, 1973) or by current-day strength models of interference effects (for a review, see M. C. Anderson & Neely, 1996). Rather, classic interference theory holds that PI involves only response competition, whereas RI involves response competition and the unlearning of

Larry L. Jacoby, Anthony J. Bishara, Sandra Hessels, and Andrea Hughes, Department of Psychology, Washington University in St. Louis.

Anthony J. Bishara is now at the Department of Psychology, Indiana University. Sandra Hessels is now at the Department of Psychology, Acadia University, Wolfville, Nova Scotia, Canada.

This research was supported by National Institute on Aging Grant AG13845. We are grateful to Carole Jacoby for assistance during all phases of the research reported here. We thank Matthew Rhodes for his input regarding earlier versions of this article.

Correspondence concerning this article should be addressed to Larry L. Jacoby, Department of Psychology, Washington University, St. Louis, MO 63130. E-mail: lljacoby@artsci.wustl.edu

List 1 responses during the learning of List 2 responses. Both response competition and unlearning are said to have their effect through an influence on associative strength. In contrast, our dual-process model postulates that recollection and accessibility bias are independent bases for responding and that both play roles in PI as well as RI.

Prior research using our dual-process approach has shown that differences in PI are sometimes completely accounted for by differences in accessibility bias (Hay & Jacoby, 1999; Jacoby et al., 2001). The experiments reported here investigated the possibility that differences in probabilistic RI can also be fully because of differences in accessibility bias, with ability to recollect being unchanged. This pattern of results would be comparable to a finding of an influence on bias, with discrimination left unchanged, in an investigation of recognition memory. We compare effects of probabilistic interference with effects produced by traditional RI conditions. Comparisons with traditional RI conditions are important for purposes of theory and are also relevant to effects of RI that have been recently prominent in the literature. As an example, the misinformation effect (e.g., Loftus & Palmer, 1974; see Ayers & Reder, 1998, for a review) conforms to an effect of RI (Jacoby, Bishara, Hessels, & Toth, 2005). Our estimation procedure allows us to assess the role of accessibility bias in producing such effects (cf. McCloskey & Zaragoza, 1985).

To investigate effects of probabilistic RI, we presented pairs of associatively related words, with each cue word paired with two different responses, and varied the probability of pairing in List 2. For example, in Phase 1 of Experiment 1, pairs of words (e.g., *knee bone*) were presented at a slow rate, and participants were told to study the pairs for a later memory test (see Figure 1). In Phase 2, pairs were presented at a fast rate, and participants were instructed to read the pairs aloud as rapidly as possible without thinking about the previous study phase when reading. They were given the cover story that the experimenter was interested in whether prior study of the word pairs would speed reading rates. In a 2/3 condition, a cue word was read twice with the response with which it had been studied (*knee bone*) and once with an alternative response (*knee bend*), whereas, in a 1/3 condition, the cue word was read only once with its studied response and twice with the alternative response. This variation in the probability of pairing constitutes a manipulation of probabilistic interference. For a direct test of memory, a cue word coupled with a fragmented version of the response (*knee b_n_*) was presented, and participants were instructed to complete the fragment with the response that was studied with the cue in Phase 1. Furthermore, they were told to produce a response to each cue word, guessing if necessary, by producing the first word that came to mind that completed the fragment and was related to the cue word.

Study List 1:	*knee bone*		
Read List 2:		1/3 condition	2/3 condition
		knee bone ×1	*knee bone* ×2
		knee bend ×2	*knee bend* ×1
Test:	*knee b_n_*		
Indirect Instructions:	Complete with the first word that comes to mind		
Direct Instructions:	Complete with the List 1 word, or failing that, guess		

Figure 1. Probabilistic retroactive interference procedure.

By the dual-process approach, when unable to recollect, participants would produce the first (i.e., most accessible) response that comes to mind, thereby showing automatic influences of memory in the form of accessibility bias. At least in part, superior recall of the target word in the 2/3 condition, compared with the 1/3 condition, is likely to result from an advantage in accessibility bias for the target that reflects its more frequent pairing with the cue in Phase 2 rather than from enhanced recollection of the target as having been studied in List 1. Accessibility bias in the 2/3 condition increases report of the studied words by increasing the likelihood that they will come to mind and be reported as a guess. We measured accessibility bias of this sort by means of an indirect test of memory. For the indirect test, participants were instructed to respond with the first word to come to mind that was related to the cue word and completed the word fragment (see Figure 1). As described next, performance on the direct test (cued recall) and performance on the indirect test were combined to determine whether probabilistic interference influenced recollection or, instead, influenced only accessibility bias.

A Dual-Process Model of Interference Effects: Dissociating Recollection and Accessibility Bias

Prior research has revealed dissociations of performance on direct and indirect tests of memory (for reviews, see Kelley & Lindsay, 1996; Richardson-Klavehn & Bjork, 1988; Roediger & McDermott, 1993). As an example, within limits, varying study time influences performance on a direct test of memory but leaves indirect test performance unchanged. Our dual-process model holds that direct test performance reflects both recollection and automatic influences of memory in the form of accessibility bias. Furthermore, for the model described below, we assume that the indirect test provides a pure measure of accessibility bias, uncontaminated by recollection, and also assume that the magnitude of accessibility bias underlying performance on a direct test of memory is equal to the level of performance on the indirect test of memory.

The assumption that the indirect test serves as a process-pure measure of accessibility bias might seem unlikely to hold. There is good reason to suspect that performance on indirect tests of memory is sometimes contaminated by intentional use of memory (e.g., Jacoby, 1991). However, the probabilistic RI procedure used in the current experiments discourages contamination of this sort. Interpolation of a list of word pairs to be read following the study list (see Figure 1) is likely to make it difficult to recollect responses studied in List 1, discouraging the use of recollection as a basis for responding for an indirect test. In addition, reading pairs whose responses fit the fragments presented at test should serve to make fragment completions readily accessible, making it unnecessary to rely on recollection. We further discouraged reliance on recollection by requiring participants to respond rapidly on the indirect test. Prior research has suggested that requiring participants to respond rapidly reduces the probability of recollection (e.g., Hay & Jacoby, 1999). To anticipate, results from our experiments provide strong support for the assumption that the indirect test does serve as a process-pure measure of accessibility bias.

For the assumption that accessibility bias on the direct test is equal in magnitude to performance on the indirect test, it is important to note that we instructed participants to produce a response to every item on the direct test, guessing if necessary. In contrast to that instruction, when investigating dissociations between performance on direct and indirect tests, it has been standard to discourage guessing on the direct test. Doing so is meant to avoid contamination of performance on the direct test by implicit memory. However, it is unlikely that effects of accessibility bias can be fully eliminated by instructing participants not to guess. Furthermore, dissociations between performance on direct and indirect tests found by instructing participants not to guess for a direct test can result from differences in response criterion rather than from differences in the types of memory underlying the two types of test (e.g., Reingold & Toth, 1996).

An instruction not to guess for the direct test produces a response criterion that is more stringent than that for the indirect test and reduces the probability of correct responding by lowering the magnitude of responding on the basis of accessibility bias. Use of a stringent criterion for the direct test can result in the finding that the probability of producing a studied item is higher for an indirect test than for a direct test of memory. However, when participants are forced to respond to each item on a direct test, equating criteria for the direct and indirect tests, our dual-process model predicts that performance on the direct test will not be lower than that on the indirect test and will be higher than on the indirect test if recollection is nonzero. Seemingly counter to this prediction, persons with amnesia performed more poorly on a cued-recall test than on an indirect test of memory (e.g., Warrington & Weiskrantz, 1970). However, their doing so might have resulted from use of a more stringent criterion for responding on the direct test because direct test responding to each item was not required. We have been unable to find results showing a disadvantage for a cued-recall test, as compared with a corresponding indirect test, when participants are required to respond to each cue for the cued-recall test.

In addition to assumptions about the equality of accessibility bias for direct and indirect tests, our estimation procedure is based on the assumption that performance on a direct test reflects the independent contributions of recollection (R) and accessibility bias (A). Given these assumptions, the probabilities (Ps) of producing an earlier studied word for a direct and indirect test are:

$$P(\text{direct}) = R + (1 - R)A, \text{ and} \tag{1}$$

$$P(\text{indirect}) = A. \tag{2}$$

Equations 1 and 2 can be combined to estimate the probability of recollection:

$$P(R) = [P(\text{direct}) - P(\text{indirect})]/[1 - P(\text{indirect})]. \tag{3}$$

Figure 2 shows the recollection/accessibility bias model in the form of a tree diagram. As described by the equations, recollection from the first studied list (which occurs with probability R) leads to a correct response. When recollection does not occur, responding depends on whether the correct response is favored by accessibility bias (A). Recollection is assumed to never occur ($R = 0$) for indirect tests.

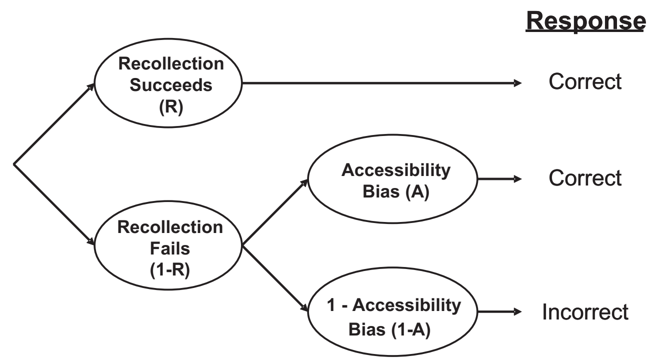

Figure 2. Multinomial processing tree diagram where correct responses can result from successful recollection (R) or from accessibility bias (A) when recollection fails ($[1 - R] \times A$). Incorrect responses are given when both recollection fails and there is a lack of an accessibility bias ($[1 - R]) \times [1 - A]$).

Again, the instruction to respond to each item on the direct test is necessary for the assumption that A for the direct test is equal to performance on the indirect test. If participants were instructed not to guess on the direct test, it would be necessary to include a threshold parameter in the model following the A parameter. The threshold parameter would represent the probability of participants having sufficient confidence in the correctness of a response to output it. A parameter of this sort is included in a model forwarded by Jacoby, Bishara, et al. (2005). An instruction to respond to each item on the direct test makes the threshold parameter equal to 1.0 and unnecessary. Also, note that the model does not include a parameter that refers to the probability of recollecting that a pair was presented in the interpolated list (List 2). A parameter of that sort is not needed. The probabilistic nature of the interference makes List 2 recollection irrelevant to responding because recollecting that a pair occurred in List 2 provides no information regarding whether or not it appeared in List 1.

The above description of the model portrays retrieval processes as being serial, with an attempt at recollection always preceding reliance on accessibility bias. However, the model would be the same if it were assumed that recollection and reliance on accessibility bias are parallel processes. We have adopted the serial description only because we believe most readers will find it easier to understand.

These estimation equations are the same as those of the one-high-threshold model that has been used to separate the effects of discriminability and response bias on recognition-memory performance (e.g., Snodgrass & Corwin, 1988). There, the goal is to separate true memory from guessing. In contrast, our dual-process model identifies true memory with recollection and holds that guessing (accessibility bias) is informed by automatic influences of memory.

To measure recollection, the difference between performance on direct and indirect tests is weighted by being divided by $[1 - P(\text{indirect})]$. This predicts that factors that selectively influence recollection will have a larger effect on direct test performance when the probability of producing the target response on an indirect test is low, meaning that accessibility bias strongly favors a competitor (high interference). In contrast, conditions that produce high accessibility bias favoring the target response (high facilitation) leave little room for differences in recollection to have an effect.

The assumption that recollection and accessibility bias are independent has been controversial for other process-dissociation procedures used to separate the contributions of the two bases for responding (e.g., Curran & Hintzman, 1997; see the response by Jacoby & Shrout, 1997). For the current procedure, a rationale for assuming their independence appeals to a difference between recollection and accessibility bias in the level of retrieval constraint that they impose (Jacoby et al., 2001). The notion is that recollection relies on cues that are not shared by competitors, whereas accessibility bias reflects a less constrained level of retrieval. For recollection, retrieval is highly constrained by the participant's effortful reconstruction of the study list context along with other cues used to access memory for presentation of the target item in that context (e.g., *knee* contextualized in the target study list). When unable to recollect, participants rely on accessibility bias by using more general cues shared by other responses (e.g., *knee* contextualized in the experiment as a whole, which includes presentations of both *knee bone* and *knee bend,* as shown in Figure 1). In the General Discussion, we further describe the notion of different levels of retrieval constraint and relate the notion to theorizing by others.

An alternative to the assumption that recollection and accessibility bias are independent is to assume that there is a redundancy relation between processes, an assumption made by generate/recognize models (e.g., Bodner, Masson, & Caldwell, 2000; Curran & Hintzman, 1997; Jacoby & Hollingshead, 1990). By those models, producing an earlier studied word as a response for a direct test requires that the word be generated as a potential response and recognized as earlier studied. Associative theories of memory have included a claim that an indirect test instructing participants to produce the first associate that comes to mind differs from a direct test only in that a direct test requires a recognition process, whereas the indirect test does not do so (e.g., Humphreys, Bain, & Pike, 1989). As is described below, a generate/recognize model requires more parameters to account for results from our experiments than does the recollection/accessibility bias model. We compared fits of the recollection/accessibility bias model to data from our Experiment 3 with those of a generate/recognize model that is similar to a model proposed by Bodner et al. (2000).

The Experiments

Investigations of the relation between performance on indirect and direct tests of memory have typically sought dissociations between the effects of manipulations on the two types of test. In contrast, we used performance on indirect and direct tests to estimate parameters that represent processes underlying performance and sought selective effects on those process parameters. The strategy is analogous to that of gaining support for a model that distinguishes between discriminability and bias by showing that some manipulations selectively influence a discriminability parameter, whereas others selectively influence a bias parameter (e.g., Snodgrass & Corwin, 1988). Experiment 1 examined the possibility that manipulating probabilistic interference, in ways illustrated in Figure 1, selectively influences accessibility bias, leaving recollection unchanged. Experiment 2 manipulated study time for pairs presented in Phase 1 along with probabilistic interference. The goal of that experiment was to reveal a double dissociation. Use of the recollection/accessibility bias model was expected to reveal that manipulating study time selectively influenced the probability of recollection, whereas manipulating probabilistic interference selectively influenced accessibility bias.

The goal of Experiment 3 was to provide a strong test of the assumptions underlying our procedure of combining results from indirect and direct tests as a means of estimating recollection and accessibility bias. To do that, we sought convergence in results from two estimation procedures based on the independence assumption but differing with regard to whether an indirect test is assumed to be a process-pure measure of accessibility bias. Results from Experiment 3 also allowed the fit of our recollection/accessibility bias model to be compared with that of a generate/recognize model.

Having gained support from Experiments 1–3 for the assumptions underlying the indirect–direct test estimation procedure, in Experiment 4, we asked whether the manipulation of probabilistic interference has effects that differ from the discrete change in responses that has traditionally been used to investigate RI. Experiment 4 included the traditional experimental condition (A-B, A-D) and the traditional control condition (A-B, C-D) for investigating RI, as well as a probabilistic interference condition.

Experiments 1A and 1B

We manipulated probabilistic interference in Experiment 1 and expected to find a large effect on both indirect and direct tests of memory. However, combining results from the indirect and direct tests of memory so as to separately estimate the contributions of recollection and accessibility bias was expected to show that manipulating probabilistic interference influenced accessibility bias but left recollection unchanged. A selective effect of probabilistic interference on accessibility bias would be consistent with results found for PI (Hay & Jacoby, 1999; Jacoby et al., 2001) but would be surprising from the perspective of classic interference theory (e.g., Postman & Underwood, 1973). A lack of an effect on the recollection estimate would show that strengthening of an interfering response had no effect on the List 1 association that is used for recollection, producing neither an effect of unlearning nor one of inhibition.

For our estimation procedure to be valid, it is necessary that the indirect test of memory be uncontaminated by intentional use of memory. We discouraged use of recollection for the indirect test by requiring participants to respond prior to a relatively short deadline. In contrast, for the direct test of memory, reliance on recollection was encouraged by not allowing participants to respond until after a short delay that they were told to use to attempt recollection. Experiment 1B differed from Experiment 1A in that it further speeded responding on the indirect test and also served to generalize results across universities.

Method

Participants. Participants in Experiment 1A were 32 undergraduates from McMaster University, Hamilton, Ontario, Canada, and participants in Experiment 1B were 24 undergraduates from

Washington University in St. Louis, Missouri. All participants were tested individually and participated in exchange for credit in an undergraduate psychology course.[1]

Materials and design. Type of test (indirect vs. direct) and probabilistic interference condition (2/3 vs. 1/3) were varied within-subjects, with each of the four combinations of conditions being represented by different stimulus sets. Words were selected from the norms reported by Jacoby (1996) to produce a pool of 66 three-word sets, with each word set including one cue word (e.g., *knee*) and two responses that were associatively related to the cue word (e.g., *bone, bend*). The two responses in each set contained the same number of letters and could be used to complete the same word fragment (e.g., *b_n_*). Four groups of 15-word sets were selected to serve as critical experimental items. The remaining six sets were reserved for buffer and practice items. The four groups of word sets that served as critical items were equated for word frequency and length for both cue words and response words, as well as for the probability of completing word fragments with each of the two response words. The critical word sets were rotated across participants through each of the combinations of within-subjects conditions so that, across formats, each response word served equally often as the target response for the memory test. To avoid primacy and recency effects, we used six word sets as buffer items in both the study and training phases and as practice items in the test phases. These items remained constant across formats.

The study list, presented in Phase 1, consisted of 66 word pairs (cue word paired with a target response word; e.g., *knee bone*), including 60 critical pairs and six buffer pairs (three at the beginning of the list and three at the end).

The probabilistic interference list, presented in Phase 2, included three presentations of each of the 60 critical cue words along with one of their two responses, as well as three presentations of each of the six buffer items (nine as primacy and nine as recency buffers). This resulted in a list of 198 pairs (cue word and response word). For pairs in the 2/3 congruent condition, a cue word appeared twice paired with the word with which it was studied during Phase 1 (e.g., *knee bone*) and once paired with its alternative response (e.g., *knee bend*). For pairs in the 1/3 congruent condition, a cue word appeared only once with its earlier studied response and twice with its alternative response. The order of the presentation of pairs to be read was random, with the restriction that pairs representing the different conditions be evenly distributed across each third of the list.

The final phase of the experiment involved two separate tests: an indirect and a direct memory test. Two test lists were constructed such that each included 30 pairs of cue words and word fragments—15 pairs in the 2/3 condition and 15 pairs in the 1/3 condition. Each test was preceded by a practice test consisting of four pairs that had served as buffers. In both the direct and the indirect tests, order of presentation was random, with the restriction that not more than three items from the same combination of conditions could be presented consecutively.

Procedure. In the study phase (Phase 1), word pairs were presented in lowercase, white letters on a black background in the center of a computer screen, using Micro Experimental Laboratory software (Schneider, 1990). Each pair was presented for 5 s, followed by a 500-ms interstimulus interval (ISI) during which time the screen was blank. Participants were instructed to read the word pairs aloud and to try to remember them for a later test of memory. To help them remember, they were instructed to think about the association between the two words and to try to form an image of the pair of items together.

In the probabilistic interference phase (Phase 2), participants were told that some of the presented word pairs would be the same as those that they had just studied, whereas others would be new. They were given the cover story that the experimenter was interested in how prior study of the word pairs would affect reading rates and told that their reading times would be recorded. Participants were instructed to read each word pair aloud into a microphone and to try to read at a constant rate, without thinking about the previous study phase when reading. Each pair of words was presented for 1.5 s, with a 500-ms ISI.

In the test phase, participants were informed that there would be two tests, the first with speeded responding (indirect) and the second with delayed responding (direct). For the indirect test (presented first for all participants), participants were presented with a cue word on the left and a word fragment on the right (e.g., *knee b_n_*). They were informed that all word fragments could be completed with two possible responses that had been seen earlier. Participants were instructed to complete the fragment as quickly as possible with the first word that came to mind and were warned that they would have a very brief amount of time in which to respond. Prior to beginning the actual test, participants completed four practice items. In Experiment 1A, each test item was presented for 1,650 ms, followed by a 1-s delay prior to presentation of the next test item. Participants were instructed to respond while the test item was on the screen, but responses were counted if they occurred during the delay. In Experiment 1B, participants were instructed to press the space bar as they gave their response. If they did not respond and press the space bar within 1,650 ms, the computer beeped to indicate that they were too slow. They were instructed to try to respond quickly enough to avoid ever hearing a beep. After the space bar had been pressed, there was a 500-ms delay prior to presentation of the next test item.

For the direct test, participants were again presented with a cue word on the left and a word fragment on the right. In contrast to the indirect test, they were instructed to complete the fragment with the word that was paired with the cue in the original study list (Phase 1). They were informed that there were two possible completion words but that only one had been presented in the study list. If they could not recall the word they had studied with the cue in the first list, they were to complete the fragment with the first word that came to mind that fit the fragment and was associatively related to the cue word. Participants completed four practice trials before beginning the actual experimental trials. Each test item was presented on the screen for 5 s before a string of asterisks (*******) appeared below the test item, signaling the participant to respond. They were instructed that they were to use the interval prior to presentation of the asterisks to recollect the Phase 1 response and that once the asterisks appeared, they were to respond immediately. Following the onset of the asterisks, participants had 1 s to respond. The experimenter entered the

[1] In this and the other experiments reported in this article, participants ranged between 18 and 23 years of age, and female participants outnumbered male participants approximately 2 to 1. For the majority of participants, credit was given for an introductory psychology course.

participant's response, and the next trial was presented after a 1-s ISI.

Computing estimates of recollection and accessibility bias. Estimates were computed both at the level of individual participants and at the level of groups. For individual participants, Equations 1–3 were used to compute estimates of recollection (R) and accessibility bias (A). Those estimates were then analyzed to separately examine the effects of interference condition on R and A. Analyzing data to obtain estimates at the level of participants is analogous to analyzing recognition-memory results by means of a single, high-threshold model to obtain estimates of discriminability (true memory) and response bias. For the recollection/accessibility bias model, performance on the indirect test serves the same role as do false alarms for the recognition-memory analysis.

A multinomial analysis was used to test the fit of the recollection/accessibility bias model to the results and to obtain estimates of R and A at the level of groups. Multinomial analyses were performed with the Microsoft Excel Solver function (Dodson, Prinzmetal, & Shimamura, 1998). Multiple random starting parameters were used so as to assure convergence on best fitting parameters and the smallest G^2, the indicator of goodness of fit. Multinomial models are said to fit the data when G^2 is below the critical value derived from the chi-square distribution. Alpha was set to .05. With this alpha, power always exceeded .999 for the detection of medium effects ($w = .3$) and .93 for the detection of small effects ($w = .1$; Cohen, 1977; Faul & Erdfelder, 1992).

To test for a significant difference between two or more parameter estimates, we compared the fitted model with a nested model that constrained the parameters of interest to be equal. The difference in G^2 between the fitted and nested models was tested against a chi-square distribution with one degree of freedom per parameter constraint. G^2 is reported with p values for nested model comparisons so as to distinguish them from general model fitting.

Multinomial model analyses were conducted on data collapsed across participants, but the results were the same when parameters were computed for each individual participant and analyzed with analyses of variance (ANOVAs). Multinomial analyses have the advantage of providing a measure of goodness of fit, which is important for the comparison of models, whereas analyses of parameters computed for individual participants have the advantage of being more easily understood by those who are less familiar with multinomial models. We report results from each of the two types of analysis and point out the parallels in results.

Results and Discussion

The data from Experiments 1A and 1B were subjected to a mixed ANOVA with experiment as a between-participants variable. It was found that neither the main effect nor any interaction effect of experiment was significant (all Fs < 2.20, all ps $> .14$). Consequently, the analyses below were conducted on the data collapsed across Experiments 1A and 1B.

Collapsed across Experiments 1A and 1B (see Table 1), recall of target items was higher for the 2/3 than for the 1/3 condition, $F(1, 55) = 118.17$, $p < .001$, $\eta_p^2 = .68$, and was higher for the direct test than for the indirect test, $F(1, 55) = 153.50$, $p < .001$, $\eta_p^2 = .74$. The difference between direct and indirect test performance was significant in both the 1/3 condition, $t(55) = 11.00$, $p < .001$, and the 2/3 condition, $t(55) = 8.40$, $p < .001$, but was larger in the 1/3 condition, as evidenced by a significant Test × Interference Condition interaction, $F(1, 55) = 19.53$, $p < .001$, $\eta_p^2 = .26$. The larger difference between performance on indirect and direct tests for the 1/3 condition was predicted by the recollection/accessibility bias model, which holds that recollection has a larger impact on performance when the probability of producing a target item because of accessibility bias is low.

Table 1
Probability of Correct Recall as a Function of Test Condition in Experiment 1

	Test condition	
Probabalistic interference condition	Direct	Indirect
1/3	.77 (.13)	.51 (.11)
2/3	.86 (.09)	.71 (.10)

Note. Values in parentheses indicate standard deviations.

These results indicate that manipulating probabilistic interference had a large effect on cued-recall performance, measured by the direct test of memory. As described next, combining results from the direct and indirect tests by means of the recollection/accessibility bias model revealed that the effects of probabilistic interference were fully because of an influence on accessibility bias, with recollection being left unchanged.

Participant-level parameter estimates. Estimates of R and A parameters were computed for individual participants. R estimates were not significantly affected by probabilistic interference (for the 2/3 and 1/3 conditions, Ms = .45 and .50, SDs = .44 and .31, respectively), $t(55) = 0.90$, $p = .37$. In contrast, A estimates, obtained from performance on the indirect test, were significantly higher in the 2/3 condition ($M = .71$, $SD = .10$) than in the 1/3 condition ($M = .51$, $SD = .11$), $t(55) = 10.47$, $p < .001$, $\eta^2 = .67$.

Group multinomial results. R was set to be 0 for the indirect test and allowed to vary for the direct test. A was expected to vary as a function of proportion congruent, and thus, A was estimated separately for the 2/3 and 1/3 conditions.

A simple three-parameter model provided a near perfect fit to these data, $G^2(1) = 0.04$, well below the critical value of 3.84. Estimated parameter values are shown in Table 2.

Even when R was allowed to vary, it was not significantly different between the 2/3 condition ($R = .52$) and the 1/3 condition ($R = .53$), $G^2(1) = 0.04$, $p = .85$. A nested model comparison revealed that the value of A was significantly higher in the 2/3 condition than in the 1/3 condition, $G^2(1) = 101.27$, $p < .001$. The lack of an effect of allowing R to vary across interference conditions, along with the significant effect of interference condition on A, corresponds to the absence of a significant effect on R, along with the significant effect on A, found in analyses of estimates gained from individual participants. Values of A were not significantly different from the probability correct on the indirect test, $G^2(2) = 0.05$, $p = .99$. A difference could have been observed because the multinomial model finds values of A that simultaneously produce the best fit for performance on indirect and direct tests, making it possible for estimates of A to differ from performance on the indirect test.

Although results from analyses of estimates obtained at the level of individual participants agreed with results from the multinomial

Table 2
Estimated Multinomial Model Parameter Values in Experiments 1 and 2

Parameter	Experiment 1	Experiment 2
R_{direct}	.53	
$R_{\text{direct-6-s}}$		.46
$R_{\text{direct-2-s}}$		.14
$A_{2/3}$	.71	.71
$A_{1/3}$	.51	.50

analyses, there was disagreement among the two types of analyses in estimates of R. Estimates of R obtained for individual participants were .50 and .45 for the 1/3 and 2/3 conditions, respectively, whereas the estimate of R obtained at the group level was .53. The smaller estimates of R calculated for individuals resulted from some participants producing a negative estimate of R because of near ceiling performance on the indirect test of memory in combination with slightly lower performance on the direct test of memory. Just as is the case for other variants of the process-dissociation procedure (e.g., Jacoby, Begg, & Toth, 1997), floor and ceiling effects can result in inaccurate estimates of R. The impact of extreme scores is reduced by gaining estimates at the levels of groups as done by the multinomial analysis. Estimates gained from groups have higher reliability than do estimates gained from individual participants.

Experiment 2

A standard finding in traditional investigations of RI is that the observed amount of RI decreases with increases in the degree of original learning (e.g., McGeoch, 1929). In Experiment 2, we manipulated degree of original learning by presenting pairs for either 2 s or 6 s for study in List 1 and also manipulated probabilistic interference. We expected the manipulation of interference to have a larger effect when degree of original learning was low rather than high (2-s vs. 6-s study). Furthermore, this difference in resistance to interference was expected to be fully because of recollection being higher in the 6-s study condition. That is, the manipulation of study time was expected to influence recollection while leaving accessibility bias unchanged, whereas the manipulation of probabilistic interference was expected to produce an opposite dissociation by influencing accessibility bias and leaving recollection unchanged. Finding a double dissociation of this sort would provide support for our assumption that recollection and accessibility bias are independent bases for responding (e.g., Hay & Jacoby, 1999). Also, a selective effect of manipulating study time on recollection would be consistent with results showing that increases in study time can enhance direct test performance without influencing indirect test performance (for a review, see Roediger & McDermott, 1993).

Results from Experiment 1 indicated that manipulating probabilistic interference selectively influenced accessibility bias. A concern for interpreting those results is that type of test was manipulated within-participants, with the indirect test of memory always preceding the direct test. This created the possibility that the shorter retention interval for the indirect test, compared with the direct test, distorted the results. To guard against this possibility, we manipulated type of test between participants in Experiment 2.

A final difference between Experiments 1 and 2 is that the number of presentations of pairs in the interpolated list was doubled in Experiment 2. That is, whereas, in the 2/3 condition in Experiment 1, a cue appeared with its earlier studied response twice and its alternative response once, in Experiment 2, a cue appeared with its earlier studied response four times and its alternative response twice. Similarly, for the 1/3 condition, the numbers of presentations of the earlier studied and alternate responses were two and four in Experiment 2. This change in procedure was meant to generalize the finding that recollection was uninfluenced by manipulating accessibility bias across a larger number of presentations of interfering responses.

Method

Participants. Forty-eight McMaster University undergraduates participated in the experiment in return for credit in an undergraduate psychology course. Participants were randomly assigned to either the direct or indirect test condition (24 in each group). Participants were tested individually.

Materials and design. Materials were the same as in Experiment 1. The study list was divided into two lists, one list for each of the two study durations. Across participants, each response word appeared equally often in each duration condition. For each participant, there was only one test list, which consisted of 60 pairs.

Procedure. In the study phase, 30 pairs were presented for 2 s each, with 500-ms ISIs. Following that, participants were told they would again study word pairs, this time presented for a longer duration. Thirty pairs were presented for 6 s each, with 500-ms ISIs. The study phase for the indirect test condition was the same as for the direct test, except that participants were instructed to remember the word pairs for a reading test rather than for a memory test. For the probabilistic interference phase, there were two blocks of word pairs, each of which was identical to the interpolated list used in Experiment 1. Thus, there were six presentations of each cue word, with the target to alternate response ratio maintained from Experiment 1 (e.g., 4:2 instead of 2:1 in the 2/3 condition). For the direct test, the response signal appeared 7 s after presentation of the test item. Once the signal appeared, participants had 2 s in which to respond. For the indirect test, participants were given a maximum of 5 s to respond but were instructed to complete the fragment as quickly as possible.

Results and Discussion

Correct recall (see Table 3) was greater for the 2/3 than for the 1/3 condition, $F(1, 46) = 94.41$, $p < .001$, $\eta_p^2 = .67$; greater for the 6-s than for the 2-s study duration, $F(1, 46) = 13.73$, $p < .001$, $\eta_p^2 = .23$; and greater for the direct than for the indirect test, $F(1, 46) = 17.30$, $p < .001$, $\eta_p^2 = .27$. The Study Duration × Test Type interaction was significant, $F(1, 46) = 7.67$, $p < .01$, $\eta_p^2 = .14$. Simple comparisons showed that study duration affected direct test performance, $t(23) = 4.83$, $p < .001$, but not indirect test performance, $t(23) = 0.63$, $p = .54$. There was also a significant Study Duration × Probabilistic Interference interaction, $F(1, 46) = 9.79$,

Table 3
Probability of Correct Recall as a Function of Test Condition in Experiment 2

Probabilistic interference condition	Test condition Direct	Indirect
$1/3_{2\text{-s}}$	.55 (.20)	.48 (.15)
$2/3_{2\text{-s}}$	.77 (.12)	.71 (.14)
$1/3_{6\text{-s}}$	.74 (.16)	.52 (.12)
$2/3_{6\text{-s}}$	.83 (.11)	.70 (.14)

Note. Values in parentheses indicate standard deviations.

$p < .005$, $\eta_p^2 = .18$. The difference between direct and indirect test performance was larger in the 1/3 condition but was significant in both the 1/3 condition, $t(46) = 3.89$, $p < .001$, and the 2/3 condition, $t(46) = 3.06$, $p < .01$.

Group multinomial results. Parameters could not be computed at the level of individual participants because type of test was manipulated between participants. For the multinomial analysis, separate *R* parameters were used for the two study durations because study time was expected to selectively affect recollection. Estimated parameter values are shown in Table 2. A four-parameter model provided an excellent fit, $G^2(4) = 3.29$, critical value = 9.49. The *R* parameter was significantly higher in the 6-s condition than in the 2-s condition, $G^2(1) = 30.77$, $p < .00001$. The *A* parameter was significantly higher in the 2/3 condition than in the 1/3 condition, $G^2(1) = 109.99$, $p < .00001$, and values of *A* were not significantly different from indirect test performance averaged across study times, $G^2(2) = 0.03$, $p = .99$. Even when allowed to vary, *R* was unaffected by the manipulation of probabilistic interference, $G^2(2) = 1.81$, $p = .41$, and *A* was unaffected by study time, $G^2(2) = 2.98$, $p = .23$.

In sum, probabilistic interference influenced accessibility bias but left recollection unchanged, whereas manipulating study time influenced recollection but left accessibility bias unchanged. Comparing the group multinomial results of Experiments 1 and 2, estimates of accessibility bias remained invariant across a wide range of levels of recollection (.14–.53). Indeed, the similarity in estimates of accessibility bias across experiments is striking (see Table 1). For both the 2/3 and the 1/3 conditions, estimates of accessibility bias are near identical across the two experiments although type of test was manipulated within participants in Experiment 1 and between participants in Experiment 2. A potentially more important difference between the experiments is that interpolated pairs were read twice as often in Experiment 2, as compared with Experiment 1, although the ratios of presentations of earlier studied and alternative responses (2/3 vs. 1/3) were the same for the two experiments. The similarity in estimates of accessibility bias across experiments suggests that it is the ratio of presentations of the two responses, rather than their number of presentations, that is important for probabilistic RI. Of course, more research is needed to further establish this invariance of accessibility bias across number of presentations and to explore its limits.

Are recollection and accessibility bias independent? An assumption underlying the recollection/accessibility bias model is that recollection and accessibility bias are independent bases for responding. This independence assumption is supported by the finding that manipulating study time selectively influenced recollection, whereas manipulating probabilistic interference selectively influenced accessibility bias. However, the independence assumption has been controversial for other procedures used to dissociate automatic and controlled processes, particularly the inclusion–exclusion procedure (e.g., Curran & Hintzman, 1997; Humphreys et al., 2003). Curran and Hintzman (1997) argued that a positive correlation between recollection and accessibility bias would result in the underestimation of accessibility bias. We describe the inclusion–exclusion procedure, along with the bases for Curran and Hintzman's arguments, when introducing Experiment 3.

Experiment 3

Experiment 3 was designed to provide a stringent test of the assumptions underlying the recollection/accessibility bias model. To gain support for those assumptions, we manipulated probabilistic RI and sought convergence in estimates of recollection and accessibility bias across two different estimation procedures. The indirect–direct test procedure for gaining estimates was the same as used in Experiments 1 and 2: Performance on the indirect and direct tests was combined by means of Equations 1–3 to gain estimates of recollection and accessibility bias.

An inclusion–exclusion procedure was used as a second means of gaining estimates (e.g., Jacoby, Toth, & Yonelinas, 1993). Instructions for an inclusion test were the same as those for the direct test of memory. For both, participants were instructed to complete fragments with earlier studied words, guessing if necessary. In contrast, for the exclusion test, participants were instructed to complete fragments with words that were not studied in List 1. For example, they were told that if they had studied the pair *knee bone* in List 1, they should not use the word *bone* to complete the corresponding test fragment (*knee b_n_*).

For the inclusion–exclusion procedure, the equations for the probability of producing an earlier studied word as a completion are:

$$P(\text{inclusion}) = R + (1 - R)A, \tag{4}$$

$$P(\text{exclusion}) = (1 - R)A, \tag{5}$$

$$P(R) = P(\text{inclusion}) - P(\text{exclusion}), \text{ and} \tag{6}$$

$$P(A) = P(\text{exclusion})/(1 - R). \tag{7}$$

The rationale underlying the exclusion equations is that participants will mistakenly produce an earlier studied word as a completion only if they do not recollect that the word was earlier studied, $P(1 - R)$, and the word comes to mind because of accessibility bias (*A*). By the inclusion–exclusion procedure, cognitive control (recollection) is measured as the probability of producing the target response when one is trying to do so (inclusion condition) minus the probability of mistakenly producing the response when one is trying not to do so (exclusion condition).

The two estimation procedures share the assumption that recollection and accessibility bias are *independent* bases for responding and also share the assumption that accessibility bias (*A*) is equal across test conditions. However, the estimation procedures differ in that the indirect–direct test procedure assumes that an indirect test serves as a process-pure measure of accessibility bias, whereas the inclusion–exclusion procedure replaces that process-pure assumption with an assumption that recollection is the same for inclusion and exclusion tests. The equations for the inclusion–exclusion procedure correspond to a two-high threshold model whereas those for the direct-indirect test procedure correspond to a one-high threshold model (e.g., Snodgrass & Corwin, 1988).

Despite the difference in assumptions underlying the two estimation procedures, we expected estimates from the two procedures to converge. For both procedures, we expected a manipulation of probabilistic interference to produce an effect on accessibility bias but to leave recollection unchanged. Furthermore, we expected the estimates of accessibility bias gained by means of the inclusion–exclusion procedure to be the same as the probabilities of producing an earlier studied word on an indirect test, showing convergence of the procedures on estimates of accessibility bias. We also expected the estimates of recollection to be the same across estimation procedures.

Returning to Curran and Hintzman's (1997) arguments against the independence assumption, they argued that a positive correlation between recollection and accessibility bias would result in the *underestimation* of accessibility bias. This underestimation would arise because, by the model used for the inclusion–exclusion procedure (Equations 4–7), A can be measured only when recollection has failed. Recollection is more likely to fail for hard items, and so, estimates of A disproportionately come from hard items. Hard items also have lower A if R is positively correlated with A, leading to a systematic underestimation of A when estimates are aggregated across items. Similarly, a positive correlation of R and A at the level of participants would result in underestimation of A when estimates are aggregated across participants because low-performing participants would more often fail to recollect, and so, events involving A would disproportionately come from low-performing participants.

In contrast to the inclusion–exclusion procedure, a positive correlation between R and A at the level of participants or items would *not* result in underestimation of A measured by the indirect–direct test procedure. For the indirect–direct test estimation procedure, measurement of A is done by means of an indirect test and so does not rely on failure of recollection. Recollection is assumed to be zero for the indirect test. The main worry for the indirect–direct test procedure is that performance on the indirect test might be contaminated by recollection rather than being process pure. Such contamination would result in performance on the indirect test *overestimating* accessibility bias.

Convergence of estimates of recollection and accessibility bias across the estimation procedures would provide strong support for assumptions underlying the estimation procedures. Curran and Hintzman (1997) argued that violation of the independence assumption results in the inclusion–exclusion procedure *underestimating A*. Violation of the assumption made for the indirect–direct test procedure that the indirect test is process pure would inflate indirect test performance and thereby further increase the difference between performance on the indirect test and A estimated by the inclusion–exclusion procedure. Consequently, finding agreement between estimates of A gained from the two estimation procedures would provide support for both the independence assumption and the assumption that the indirect test provides a process-pure measure of accessibility bias.

Method

Participants. Sixty-four McMaster University undergraduates participated in the experiment in return for credit in an undergraduate psychology course. Participants were randomly assigned to either the exclusion–inclusion or the direct–indirect test condition (32 in each group). Participants were tested individually.

Materials, design, and procedure. The materials were the same as those used in Experiment 2. In Phase 1, study pairs were presented for 3 s each. The interpolated list of pairs that was presented to be read in Phase 2 was the same as in Experiment 2. That is, for the 2/3 condition, the ratio of presentations of the earlier studied and alternate responses was 4:2, and for the 1/3 condition, the ratio was 2:4.

In the test phase, participants in the exclusion–inclusion condition first completed the exclusion test. They were told that they would be presented with a cue word on the left and a word fragment on the right and were instructed to complete the fragment with a word that was associatively related to the cue word but had not been presented in the study phase. If they could not recall the earlier studied response, then they were to complete the fragment with the first word that came to mind that fit the fragment and was associatively related to the cue word. Participants were instructed to respond immediately after the appearance of a string of asterisks below the test item. The asterisks appeared 7 s after presentation of the test item. Once the asterisks appeared, participants had 2 s to respond. The exclusion–inclusion participants then completed the inclusion test. They were given the same instructions as for the exclusion test, except that this time they were to complete the fragment with the word that was studied. Participants in the direct–indirect condition completed the indirect test and then the direct test as in Experiment 1. For both the exclusion–inclusion condition and the indirect–direct condition, other details of the method were the same as for Experiment 1.

Results and Discussion

Probabilities of correct recall (see Table 4) were compared with a mixed-factor ANOVA, with interference condition and test type as within-subjects variables and test condition (direct–indirect or inclusion–exclusion) as a between-subjects variable. The Test Type $\times$ Test Condition interaction was significant, $F(1, 62) = 20.64$, $p < .001$, $\eta_p^2 = .25$, as was the three-way Interference Condition $\times$ Test Type $\times$ Test Condition interaction, $F(1, 62) = 5.17$, $p < .05$, $\eta_p^2 = .08$.

Table 4
Probability of Correct Recall as a Function of Estimation Procedure by Test Condition in Experiment 3

Probabilistic interference condition	Estimation procedure: Direct–indirect test condition		Estimation procedure: Inclusion–exclusion test condition	
	Direct	Indirect	Inclusion	Exclusion
1/3	.70 (.19)	.47 (.14)	.67 (.16)	.32 (.14)
2/3	.82 (.13)	.69 (.13)	.82 (.09)	.43 (.16)

Note. Values in parentheses indicate standard deviations.

Participant-level parameter estimates. Estimates of R were submitted to a 2 (interference condition) × 2 (test procedure: direct–indirect vs. inclusion–exclusion) mixed ANOVA. There were no significant main or interaction effects (all Fs < 1, ps > .49). When the same ANOVA was applied to estimates of A, there was only a main effect of interference condition, $F(1, 62) = 56.32$, $p < .001$, $\eta_p^2 = .48$. Test procedure did not have a significant main effect, nor did it significantly interact with interference conditions (both Fs < 1, ps > .61).

Estimates of A for the 1/3 and 2/3 interference conditions gained by means of the inclusion–exclusion procedure (Ms = .49 and .69, SDs = .19 and .15) were near identical to performance for the corresponding conditions on the indirect test (Ms = .48 and .68, SDs = .13 and .11).[2] This convergence in results provides strong support for the assumption that recollection and accessibility bias serve as independent bases for responding and the assumption that the indirect test serves as a process-pure measure of accessibility bias.

Group multinomial results. To test for convergence among different measurement methods, we constrained the R parameter to be the same for direct, inclusion, and exclusion tests and set R to 0 for the indirect test. The A parameter was expected to be influenced by interference condition but not by test type, and so, only two A parameters were used. Estimated values are shown in Table 5. This simple three-parameter model provided a very good fit, $G^2(5) = 2.23$, critical value = 11.07. The A parameter was significantly higher in the 2/3 condition than in the 1/3 condition, $G^2(1) = 103.49$, $p < .00001$, and values of A were not significantly different from indirect test performance, $G^2(2) = 1.24$, $p = .54$. Even when allowed to vary, R did not significantly differ across direct, inclusion, and exclusion tests, $G^2(2) = 1.31$, $p = .52$, and neither did A, $G^2(4) = 2.20$, $p = .70$. Also, R did not differ across the 2/3 and 1/3 conditions, $G^2(1) = 0.18$, $p = .67$.

To further examine whether the indirect test provided a process-pure measure of A, we tested a model where R was free to vary for the indirect test. The model converged on a stable value of .00 for R_{indirect}. Furthermore, when R_{indirect} was constrained to various values, the model had to be rejected whenever R_{indirect} exceeded .05. These results indicate that contamination of the indirect test was nonexistent or, at most, negligible.

Table 5
Estimated Multinomial Model Parameter Values for Recollection/Accessibility Bias (R/A) Model and Generate/Recognize (G/R) Model in Experiment 3

	Model	
Parameter	R/A	G/R
R	.38	.50
$A_{2/3}$	.70	.69
$A_{1/3}$	.49	.47
T		1.00
G		.40
df	5	3
G^2	2.23	3.16
Critical value	11.07	7.81
AIC	8.23	13.16
BIC	26.99	44.43

Note. The R parameter differs for the two models. For the R/A model, R refers to recollection, and for the G/R model, R refers to recognition. AIC = Akaike information criterion; BIC = Bayesian information criterion.

One might argue that the similarity of the equations for the two estimation procedures is sufficiently high to make the convergence of results uninformative. Against that criticism, we swapped the indirect equations with exclusion equations and then attempted to fit the three-parameter model (R, $A_{2/3}$, and $A_{1/3}$). The model no longer fit the data, $G^2(5) = 298.40$, critical value = 11.10. The extremely poor fit when equations were swapped shows that the differences among assumptions for the two estimation procedures are important.

The convergence in estimates of A across the two estimation procedures provides strong support for the assumption that recollection and accessibility bias serve as independent bases for responding and the assumption that the indirect test serves as a process-pure measure of accessibility bias. By Curran and Hintzman's (1997) arguments, a violation of the independence assumption would produce underestimation of A by the inclusion–exclusion procedure that would result in A being lower than performance on the indirect test. Contamination of the indirect test by recollection would also result in A being lower than performance on the indirect test. That is, violation of either assumption would produce a difference between A and indirect test performance, and violation of both assumptions would have the additive effect of producing a large difference between A and indirect test performance. The convergence also provides strong support for the assumption that the R for the inclusion test is equal to the R for the exclusion test (cf. Humphreys et al., 2003), which is made for the inclusion–exclusion means of obtaining estimates. Serious violation of any of these assumptions would be expected to prevent convergence of estimates across procedures.

Generate/recognize model. An alternative to assuming that recollection and accessibility bias serve as independent bases for responding is to argue that direct test performance relies on a generate/recognize process (e.g., Bodner et al., 2000; Curran & Hintzman, 1997). By generate/recognize models, producing an earlier studied word as a response for an indirect test requires that the word be generated as a potential response and recognized as earlier studied. If participants generate only a single response for each item on a direct test of memory, a generate/recognize model would predict that performance on the direct test would be lower than or, at most, equal to performance on an indirect test of memory. Instructing participants to produce a response for each item on the direct test would effectively remove the recognition criterion and result in performance on the direct and indirect tests being equal. To account for performance on a direct test being higher than that on an indirect test, as found in our experiments, a generate/recognize model has to include an assumption that multiple candidate responses are sometimes generated for test items on the direct test (e.g., Bodner et al., 2000; Jacoby & Hollingshead, 1990).

We compared fits of the recollection/accessibility bias model to data from Experiment 3 with those of a generate/recognize model

[2] One participant had to be excluded from the participant parameter estimates because of ceiling performance on the indirect test, which leads to an infinite R parameter.

similar to the model proposed by Bodner et al. (2000). In this model, A represents the probability of automatically generating the earlier studied word as a first candidate response, and Rg represents the probability of recognizing a generated word as studied in List 1. T represents the probability of responding with an earlier studied word even when it was not recognized. Finally, G represents the probability of generating an earlier studied word as a second or later candidate response for a direct test of memory.

By this model, a correct response for a direct test (inclusion test) can result from the response coming to mind because of A and being recognized as studied in List 1 (ARg) or because of its being given as a response when not recognized ($A[1 - Rg]T$). If these were the only two bases for correct responding, performance on a direct test could not exceed that on an indirect test of memory. However, G serves as an alternative to A as a means of generating a correct response that can be produced because it is either recognized ($[1 - A]GRg$) or not recognized ($[1 - A]G[1 - Rg]T$):

$$P(\text{direct}) = ARg + A(1 - Rg)T + (1 - A)GRg + (1 - A)G(1 - Rg)T. \quad (8)$$

For an exclusion test, a List 1 response would be mistakenly produced only if it was generated because of either A or G and given as a response without being recognized ($[1 - Rg]T$):

$$P(\text{exclusion}) = A(1 - Rg)T + (1 - A)G(1 - Rg)T. \quad (9)$$

Just as done for the recollection/accessibility bias model, we assumed that the indirect test provides a process-pure measure of A.

We attempted to fit several different versions of the generate/recognize model. The version that fit with the fewest free parameters had five free parameters: $A_{1/3}$, $A_{2/3}$, Rg, T, and G. These parameters are similar to the ones used for the recollection/accessibility bias model, with Rg being substituted for R and with T and G being added. As shown in Table 5, this model produced a good fit to the data. The A parameter was significantly higher in the 2/3 condition than in the 1/3 condition, $G^2(1) = 101.70$, $p < .001$. Even when allowed to vary, neither the Rg nor the T parameter differed across the 2/3 and 1/3 conditions, both $G^2(1) < 2.22$, $ps > .13$. The G parameter, when allowed to vary, did not significantly differ across direct inclusion versus exclusion instructions, $G^2(1) = 2.19$, $p = .14$.

Note that the T parameter was estimated to be 1.00. When this happens, the generate/recognize model becomes algebraically equivalent to the recollection/accessibility bias model for the direct, indirect, and inclusion conditions. As can be seen by examining Equation 8, when $T = 1$, the Rg parameter drops out of the equation, and the equation simplifies to become $A + (1 - A)G$, which equals $G + (1 - G)A$. The G parameter mimics recollection, whereas the A parameters act the same way across models, with the result that the generate/recognize model becomes the recollection/accessibility bias model. As shown in Table 5, the A and G parameter estimates from the generate/recognize model are quite close to A and R from the recollection/accessibility bias model. The only reason they are not identical is that the equations differ in the exclusion condition even when $T = 1$.

The three-parameter recollection/accessibility bias model that fits is not nested within this five-parameter generate/recognize model that also fits. Consequently, the two models cannot be compared by using G^2 alone. We used a combination of two complexity-adjusted fit indicators known to have two different tendencies. The first is the Akaike information criterion (AIC; Akaike, 1974), which tends to favor more complex models, especially when the sample size is large. The second adjustment is the Bayesian information criterion (BIC; Schwarz, 1978), which tends to favor simpler models. Because they have opposing tendencies, when both complexity adjustments agree, there is good evidence that one model is performing better than the other (Burnham & Anderson, 2004; Kuha, 2004). AIC involves adding a penalty to G^2 of $2k$, where k is the number of parameters in the model. BIC involves adding a penalty of $k \times \ln(n)$, where n is the total number of observations. For both AIC and BIC, smaller values mean better model performance. As shown in Table 5, both indicators show that the recollection/accessibility bias model should be preferred.

Experiment 1 did not allow enough degrees of freedom to test the generate/recognize model, which necessarily has more parameters than does the recollection/accessibility bias model. However, to be preferred, the generate/recognize model would have to justify its additional parameters by showing that its fit produced a G^2 much smaller than that produced by the recollection/accessibility bias model. The fit of the recollection/accessibility bias model, $G^2(1) = 0.04$, was sufficiently good that it would be impossible to significantly improve the fit by adding parameters. We also examined the fit of the generate/recognize model to the data from Experiment 2. However, with only direct and indirect test data available in that experiment, the parameter used to represent generation of additional responses cannot be statistically disentangled from the recognition-memory parameter because the two parameters trade off.

To summarize, Experiments 1–3 provided strong support for assumptions underlying the recollection/accessibility bias model. Results from each of the experiments showed that manipulating probabilistic interference selectively influenced accessibility bias, having no effect on recollection. In contrast, manipulating study time (Experiment 2) selectively influenced recollection, having no effect on accessibility bias. Results of Experiment 3 provided further support for assumptions underlying the recollection/accessibility bias model by showing convergence of parameter estimates across two estimation procedures. Violation of assumptions would have produced estimates of A that differed from performance on an indirect test of memory. The results of Experiment 3 can be fit by the generate/recognize model. However, the generate/recognize model fits the results by becoming the recollection/accessibility bias model for the direct (inclusion) test, and its additional parameters are unjustified by the results.

Experiment 4

Having gained support for its underlying assumptions, in Experiment 4, we used the indirect–direct test estimation procedure to compare effects in a probabilistic interference condition with those in the traditional pure interference condition (A-B, A-D) and the traditional control condition (A-B, C-D). Comparisons with the pure interference condition are important for determining the effect of intermixing presentations of A-B (the List 1 pair) with presentations of A-D in the interpolated list, as done to produce probabilistic interference. It is possible that presenting the List 1 pair in the interpolated list has an effect on recollection as well as its influence on accessibility bias. An effect on recollection could arise from participants being reminded of the List 1 presentation of the pair when the pair is presented again in the interpolated list. This reminding would serve to reinforce the List 1 association and enhance recollection in the same way as would repeated presentation of the pair in List 1. Against this possibility, List 1 and the interpolated list were well differentiated in our experiments. List 1 was presented at a slow rate, with participants being instructed to study the words for a subsequent test. In contrast, List 2, the interpolated list, was presented at a fast rate under the guise of a test of reading speed, and participants were instructed to avoid thinking back to List 1 while engaged in the reading task. Because remindings of List 1 were discouraged by list differentiation and by instructions, we expected recollection for the probabilistic interference condition not to differ from that in a pure interference condition, which would not provide an opportunity for reminding. Rather, we predicted that the probabilistic interference and pure interference conditions would differ only in that accessibility bias would be lower in the pure interference condition as compared with the probabilistic interference condition. Presenting the List 1 pair in the interpolated list for the probabilistic interference condition was expected to increase accessibility bias favoring the List 1 response.

A comparison of effects in the interference conditions with those in the traditional RI control condition allows one to examine the possibility that interference has no influence on recollection but, instead, fully has its effects by means of an influence on accessibility bias. Traditional interference theory (e.g., Postman & Underwood, 1973) holds that RI results from unlearning as well as from response competition. Identifying unlearning produced by interference with an influence on recollection would predict that recollection in the interference conditions would be lower than that in the traditional control condition.

Method

Participants. Thirty-six Washington University in St. Louis undergraduates participated in the experiment in return for credit in an undergraduate psychology course. Participants were tested individually.

Materials and design. Six groups of 20 cue words, along with their response words plus six buffer items, were constructed, chosen mostly from the same item pool as the previous experiments and using the same criteria. As before, groups were balanced on word frequencies, probability of completing word fragments with each of the relevant response words when new, and word lengths. Word-set groups were rotated across participants through the three experimental conditions (study only, 2/4 interference, pure interference) and through the two test types (indirect, direct). As well, across participants, each response word served equally often as the target response. This resulted in 12 formats (three experimental conditions × two test types × two response targets). The Phase 1 study list consisted of 120 critical pairs. The Phase 2 interpolated list consisted of 320 pairs: four presentations of each of 80 critical items (all but the 40 study-only items). For the 2/4 interference condition, the studied target response word was presented on two trials, and the alternate response word was presented on two trials. For the pure interference condition, the alternate response word was presented on all four trials. Both test lists (indirect and direct) consisted of 60 cue–fragment pairs (20 pairs for each condition). A practice test, consisting of three cue–fragment pairs, preceded each test list.

Procedure. Presentations of stimuli and instructions for the study phase and the interference phase were identical to Experiment 3, with the exception that study pairs presented in Phase 1 were presented for 5 s each. In the indirect test phase, participants were to complete the fragment in a maximum of 1.5 s. As in Experiment 1, as participants gave their response, they were to press the space bar. If they did not respond and press the space bar within the allotted time, then the computer beeped to indicate that they were too slow. In the direct test phase, participants were to complete the fragment immediately after the response signal, which appeared 5 s after presentation of the test item. Once the signal appeared, participants had 4 s in which to respond.

Results and Discussion

Probabilities of correct recall were analyzed in a 3 (RI condition: 2/4 interference, pure interference, and study only) × 2 (test type: direct and indirect) repeated measures ANOVA. Correct recall (see Table 6) was greater for the direct than for the indirect test, $F(1, 35) = 213.14$, $p < .001$, $\eta_p^2 = .86$. There was also a significant main effect of RI condition, $F(2, 70) = 158.58$, $p < .001$, $\eta_p^2 = .82$. Paired *t* tests showed that the three RI conditions were all significantly different from one another (*p*s < .001). The Test Type × RI Condition interaction was significant, $F(2, 70) = 5.42$, $p < .01$, $\eta_p^2 = .13$. The difference between direct and indirect tests was smallest in the 2/4 condition, though it was significant in all three conditions (all *p*s < .001).

Participant-level parameter estimates. Estimates of *R* were submitted to a repeated measures ANOVA, which revealed a significant effect of RI condition, $F(2, 70) = 3.33$, $p < .05$, $\eta^2 = .09$. Post hoc paired *t* tests showed that *R* was higher in the study-only condition (.53) than in the other two conditions (*p*s < .05). The 2/4 interference and pure interference conditions (.39 and .42) were not significantly different from one another ($p = .71$). A repeated measures ANOVA revealed that *A,* measured by performance on the indirect test, varied significantly across conditions, $F(2, 70) = 104.60$, $p < .001$, $\eta^2 = .75$. Estimates of *A* in the study-only and 2/4 interference conditions did not differ from one another ($p = .20$) but were both significantly higher than *A* in the pure interference condition (*p*s < .001).

Group multinomial results. To test the possibility that recollection was the same in the 2/4 and pure interference conditions,

Table 6
Probability of Correct Recall as a Function of Test Condition in Experiment 4

	Test condition	
Probabilistic interference condition	Direct	Indirect
Study only	.80 (.13)	.57 (.11)
2/4	.74 (.13)	.55 (.14)
Interference	.61 (.15)	.32 (.13)

Note. Values in parentheses indicate standard deviations.

we constrained R to be equal for those two conditions but allowed it to be different for the study-only condition. The A parameter was allowed to be different for each of the three conditions. This model provided a near perfect fit, $G^2(1) = 0.09$, critical value $= 3.84$. R was significantly higher in the study-only condition (.53) than in the 2/4 and pure interference conditions (.42), $G^2(1) = 4.30$, $p = .04$. Even when allowed to vary, R did not differ across the 2/4 (.41) and pure interference (.43) conditions, $G^2(1) = 0.09$, $p = .76$. $A_{\text{interference}}$ (.32) was significantly lower than the other two A parameters (.56; ps $< .00001$), but $A_{\text{study-only}}$ and $A_{2/4}$ did not differ from one another, $G^2(1) = 1.11$, $p = .29$. Values of A were not significantly different from indirect test performance, $G^2(3) = 0.04$, $p = .998$.

This experiment did not provide sufficient degrees of freedom to allow the generate/recognize model to be fit to the results. However, the fit of the recollection/accessibility bias model to the results was so good that its fit would necessarily be better than that of a generate/recognize model if the larger number of parameters required by the generate/recognize model was taken into account.

Results analyzed with the recollection/accessibility bias model showed that intermixing presentations of List 1 associates with those of competitors in the interpolated list did not enhance recollection but, rather, only influenced accessibility bias as compared with the pure interference condition. The lack of an influence on recollection provides evidence against the possibility that remindings of the List 1 associate during the interpolated list served to strengthen the association upon which recollection of the List 1 response was based. Of course, if the experimental situation were changed in ways that encouraged such remindings, an influence on recollection might be obtained. Also, it is likely to be important that the manipulation of interference condition was within participants. Had the manipulation been between participants, recollection of a response as having appeared in List 2 could have served as a basis for its rejection as a List 1 response in the pure interference condition but could not have done so in the probabilistic interference condition, and that difference might be important. Intermixing presentations of List 1 pairs with List 2 presentation of competing responses, as done for the within-participant manipulation of interference condition, makes List 2 recollection uninformative for both interference conditions.

Comparison of the traditional interference and control conditions revealed that the two conditions differed both in recollection and in accessibility bias. The difference in recollection is unsurprising given the classic literature devoted to investigations of RI. The notion of accessibility bias is similar to that of response competition, and the classic literature has shown that response competition alone is insufficient to account for RI. Melton and Irwin (1940) found that overt intrusions accounted for only a portion of RI and attributed the remaining RI to what they called *Factor X.* They identified Factor X with unlearning of List 1 responses during the learning of List 2. The intrusion of List 1 responses during the learning of List 2 was said to lead to the intruding responses not being reinforced in that new context, which resulted in their being unlearned or extinguished.

By the recollection/accessibility bias model, Factor X can be identified with the reduction in recollection observed in the interference conditions as compared with the A-B, C-D control condition. However, we hesitate to attribute that reduction in recollection to unlearning or extinction of the List 1 association. In contrast to the procedure used in traditional investigations of RI, our participants were not instructed to learn the interpolated list but, rather, simply to read pairs in that list without attempting to remember them for a later test. Reading pairs in the interpolated list seems unlikely to have required unlearning or inhibition of List 1 associations. Furthermore, if it did so, unlearning should have been greater in the pure interference condition as compared with the probabilistic interference condition because competing responses were read more often in the pure interference condition. However, recollection did not differ for those conditions.

Mensink and Raaijmakers (1989) proposed a model of interference effects that is based on the Raaijmakers and Shiffrin (1981) search of associative memory (SAM) model but includes a time-dependent, contextual fluctuation process allowing it to account for classic interference effects. Contextual retrieval strength is assumed to depend on the overlap in the contextual elements encoded in a memory trace and elements active at the time of test. Mensink and Raaijmakers showed that the contextual fluctuation process allows the model to account for RI effects without assuming that interference results in the unlearning or inhibition of List 1 responses. Their model is a generate/recognize model in that a list-discrimination process follows successful recovery of a memory trace and tests whether the trace is from the to-be-recalled list.

The advantage of the recollection/accessibility bias model over the generate/recognize model in the current experiments suggests that instructions to recall List 1 responses, as compared with an indirect test, have an influence on retrieval processes rather than on postretrieval list discrimination. A possible account of the reduction in recollection produced by interference is to argue that recollection requires the effortful reinstatement of study context and that such reinstatement is easier if the cue provided for retrieval has occurred only in List 1.

Distinguishing between effects on accessibility bias and effects on recollection provides additional constraints for theorizing about RI and raises new questions. For example, as compared with the traditional control condition, does PI also influence both recollection and accessibility bias, or are there inhibitory processes, such as retrieval inhibition, that contribute only to RI (e.g., M. C. Anderson & Spellman, 1995)? In the General Discussion section, below, we further consider the relationship between RI and PI, along with the role of effortful reinstatement of context for recollection.

General Discussion

Using the recollection/accessibility bias model, results from our experiments showed that manipulating probabilistic interference influenced accessibility bias but left recollection unchanged, whereas manipulating study time (Experiment 2) produced an opposite dissociation by influencing recollection but leaving accessibility bias unchanged. Experiment 3 used an inclusion–exclusion procedure as well as the indirect–direct test procedure as means of obtaining estimates. The convergence of estimates across estimation procedures provides strong support for assumptions underlying the estimation procedures. Experiment 4 compared effects of probabilistic interference with effects in the standard conditions used to investigate RI. Results from that experiment showed that the difference between effects in a probabilistic interference condition as compared with the standard pure interference condition was fully because of a difference in accessibility bias. In contrast, the standard control condition held an advantage both in recollection and in accessibility bias over the pure interference condition.

Convergence of Results for RI and PI: Probability Matching

The patterns of results found for probabilistic RI converge with results from investigations of PI. Just as found for RI, we (e.g., Hay & Jacoby, 1999; Jacoby et al., 2001) have shown that for PI, manipulating study time selectively influences recollection, whereas manipulating probabilistic interference selectively influences accessibility bias. However, both the procedure for manipulating probabilistic interference and the means of estimating recollection and accessibility bias differ between our investigations of PI and RI.

The PI experiments used the same materials as did our RI experiments but varied probabilistic interference in a training phase that preceded presentation of the list for which memory was to be tested. Also, the number of pairings used to create probabilities in the PI experiments was much larger than that used in the RI experiments. For example, to produce a 2/3 interference condition, the cue word was presented with one response (e.g., *knee bone*) 12 times and with its alternative response (e.g., *knee bend*) six times (Hay & Jacoby, 1999, Experiment 2). Thus, the number of pairings used to create a 2/3 condition for investigating PI (12:6) was much larger than the number of pairings used to create a 2/3 condition for investigating RI (2:1 or 4:2).

The procedures for estimating recollection and accessibility bias also differed. The estimation procedure for our investigations of PI combined results from a congruent condition with results from an incongruent condition. For the congruent condition, the response presented in the study list (List 2) was the one that had been presented most often during training, causing recollection and accessibility to yield the same response. For that condition, correct recall could result either from recollection (R) of the List 2 word or, when recollection failed ($1 - R$), from reliance on the accessibility bias (A) developed during training: $P(\text{correct}|\text{congruent}) = R + (1 - R)A$. For the incongruent condition, in contrast, accessibility bias and recollection were placed in opposition by having participants study the response that was presented less often in training. For incongruent pairs, false recall (saying *bone* when *bend* was presented in List 2) would occur when participants failed to recollect the study pair ($1 - R$) and instead relied on accessibility bias (A): $P(\text{incorrect}|\text{incongruent}) = A(1 - R)$. These estimation equations are the same as the equations used by the inclusion–exclusion estimation procedure (Equations 4–7).

However, when the equations are used to combine results from congruent and incongruent conditions, the estimate of recollection is not a pure one but, rather, measures the contribution of the List 1 presentation of a pair, including its contribution to accessibility bias as well as recollection. The estimate of accessibility bias gained from the congruent–incongruent procedure measures only accessibility bias created by prior training and does not include the effect of the study presentation of a pair on accessibility bias. A striking result is that using the congruent–incongruent estimation procedure to investigate effects of PI yields estimates of accessibility bias that show probability matching (Hay & Jacoby, 1999; Jacoby et al., 2001). For example, in Hay and Jacoby's (1999) Experiment 2, the estimate of accessibility bias for a 2/3 condition was .64, which is close to .67, the training probability.

For our PI experiments, test items in a no-study condition served a function similar to that of an indirect test of memory as a measure of accessibility bias. For the no-study condition, a cue word was paired with two different responses during training, just as was done for congruent and incongruent conditions, and the memory test was also the same as for the other conditions. However, neither the cue word nor its responses were presented in List 2. Participants were warned that the test list would include cue words and fragments that did not correspond to a studied pair and were instructed to respond to such pairs by producing the first word that came to mind that fit the fragment and was associatively related to the cue word.

Results from the no-study condition showed probability matching and closely agreed with estimated accessibility bias. In Hay and Jacoby's (1999) Experiment 2, the probability of producing the typical response on no-study tests was .67, showing perfect probability matching, and estimated accessibility bias was .64. The convergence between performance on an indirect test (no-study condition) and estimates of accessibility bias gained from the congruent–incongruent estimation procedure for PI is similar to that found for RI in Experiment 3.

The direct test data from our RI experiments can be analyzed by means of the congruent–incongruent estimation procedure, with the 2/3 condition corresponding to a congruent test and the 1/3 condition corresponding to an incongruent test. Direct test data from Experiment 2 were analyzed in this way. Doing so is equivalent to using a free A parameter to represent accessibility bias in the 2/3 condition, just as before, but constraining accessibility bias in the 1/3 condition to equal $1 - A$. The data were well fit by this model, which used two free R parameters for the two study durations and one free A parameter, $G^2(1) = 1.76$, critical value = 3.84. R was significantly higher in the 6-s condition (.57) than in the 2-s condition (.32), $G^2(1) = 30.79$, $p < .001$. Even when allowed to vary, R was uninfluenced by proportion congruent, $G^2(2) = 1.76$, $p = .41$. In contrast, the A parameter (.64) was uninfluenced by study duration, $G^2(1) = 1.76$, $p = .18$.

Note that the estimate of accessibility bias for RI shows probability matching, converging with results found by Hay and Jacoby (1999) for PI. That estimate of accessibility bias is lower than the estimate of accessibility bias provided by an indirect test of RI (.71). This is to be expected because the indirect test reflects the contribution of the study presentation and the contribution of the interpolated list, whereas, for the congruent–incongruent analysis, accessibility bias measures only the contribution from the interpolated list. In the same vein, estimates of recollection gained from the congruent–incongruent analysis (.57 and .32) were larger than estimates of recollection that came from the indirect–direct test procedure (.46 and .14) because estimates from the congruent–incongruent analysis included the contribution of the study presentation to accessibility bias as well as the contribution of recollection.

In sum, results from our RI experiments converge with those from PI experiments in showing selectivity of effects of manipulations on estimates of recollection and estimates of accessibility bias. For both RI and PI, estimates of accessibility bias are found to be near identical to performance on an indirect test (a no-study condition for PI). Furthermore, measures of accessibility bias restricted to contributions from the interference phase show probability matching. This convergence of findings provides strong support for assumptions underlying the estimation procedures and is also important for theories of interference effects.

Curran and Hintzman (1997), as well as others (e.g., Bodner et al., 2000), have been critical of the independence assumption underlying the process-dissociation procedure. However, findings that probabilistic interference selectively influences accessibility bias and that manipulating study time selectively influences recollection provide support for the independence assumption, as does the finding of converging results across estimation procedures. One cannot prove that an independence assumption is true. However, the evidence to support the independence assumption for the experiments reported in the current article as well as for our PI experiments is sufficiently strong to provide a serious challenge for critics of the assumption.

A Dual-Process Model of Interference Effects

M. C. Anderson and Neely (1996, p. 249) described interference effects in terms of a ratio-rule equation (Luce, 1959), illustrating its use with an example: "p(recall *rock*, given *dog*) = Strength(*dog-rock*)/Strength(*dog-rock*) + Strength(*dog-sky*) . . . Strength(*dog-N*th item)." They noted that the ratio-rule equation can be found in the relative strength retrieval assumptions adopted by theories of interference effects (e.g., J. R. Anderson, 1983; Raaijmakers & Shiffrin, 1981). By the ratio-rule equation, as the associations between competing items become stronger, the probability of recalling the target word from a particular associative pair decreases because of an increase in the denominator.

The ratio rule provides an excellent description of effects of probabilistic interference on accessibility bias. However, it does not provide a means of accounting for the differences in performance between indirect and direct tests of memory, nor can it explain the selectivity of effects of manipulations on recollection and accessibility bias that we have observed. The dual-process model accounts for these results by proposing that recollection serves as a second basis for responding. As compared with accessibility bias, recollection relies on retrieval processes that are highly constrained by the participant's effortful reinstatement of the specified study list context along with other contextual cues used to access memory for presentation of the target item in List 1 (e.g., *knee* contextualized in the particular study list). In contrast, accessibility bias relies on more general cues that are shared by other responses (e.g., *knee* contextualized in the experiment as a whole, which includes its presentations paired with an interfering response in the interpolated list), and the ratio rule holds.

Our emphasis on the importance of reinstating context and on differences in cuing produced by test instructions agrees with theorizing by others (e.g., Humphreys, Bain, & Pike, 1989; Humphreys et al., 2003; Mensink & Raaijmakers, 1989). Our approach is also generally consistent with Nelson's PIER2 model of cued recall (Nelson, McKinney, Gee, & Janczura, 1998). However, those models postulate that a generate/recognize process plays an important role in cued-recall performance, whereas our results indicate that the recollection/accessibility bias model provides a better account of the effects of probabilistic interference than does a generate/recognize model.

Early theorizing about interference effects implicated a generate/recognize process by emphasizing the importance of list discrimination. To overcome response competition, a person's ability to identify a response as coming from List 1 or List 2 was said to be important (e.g., Winograd, 1968). Commenting on the role of list discrimination, Postman and Underwood (1973) said that "the critical factor is not the subject's ability to identify the list membership of whatever responses do occur but rather the mechanism governing the availability of alternative response repertoires for recall" (p. 24). Similarly, we argue that it is important to distinguish between context as a feature to be identified versus context as a constraint on retrieval. That is, rather than viewing the source of an event (e.g., List 1 vs. List 2) as a feature to be identified after a (nonsourced) item has come to mind, a more important use of source is to constrain the retrieval process itself so that only particular, relevant past events are brought to mind (e.g., Jacoby, Shimizu, Velanova, & Rhodes, 2005).

By the recollection/accessibility bias model, recollection relies on the effortful reinstatement of study context, and the similarity among contexts (list differentiation) is likely to influence recollection. As shown by the comparison of the standard control condition with interference conditions in Experiment 4, interference can influence recollection and accessibility bias. However, for RI and PI, we have shown that increased interference effects can result fully from an influence on accessibility bias alone. Doing so shows the importance of distinguishing between recollection and accessibility bias for theorizing about interference effects.

Distinguishing between recollection and accessibility bias is important for interpreting memory deficits. Memory deficits revealed by older adults and by frontal-lobe patients are largest in high-interference situations, and their being so has been explained as resulting from an inhibition deficit (e.g., Hasher & Zacks, 1988; Shimamura, 1995). However, by the recollection/accessibility bias model, that greater disadvantage could result from a deficit in recollection even if a deficit in inhibition did not play a role. In that vein, presenting study pairs for a short, rather than a longer, duration reduced recollection and produced a larger effect when interference was high (Experiment 2). However, the only thing special about high interference when study time was manipulated was that the lower accessibility bias provided increased opportunity for recollection to have an effect. The same may be true for effects of aging (Jacoby et al., 2001).

As an example, suppose that an experiment showed that in a low-interference condition, the probabilities of correct cued-recall for young and older adults were .92 and .84, respectively, whereas the corresponding probabilities in a high-interference condition were .68 and .36. One might conclude that older adults showed a substantial deficit only when it was necessary to inhibit a prepotent response. However, the recall probabilities for the example were generated from the recollection/accessibility bias model by setting the probability of recollection at .6 for young adults and at .2 for older adults and by setting accessibility bias at .8 for the low-interference condition and at .2 for the high-interference condition. The example illustrates that the same difference in ability to recollect can produce a small difference (low-interference condition) or a large difference (high-interference condition) in performance dependent on the level of accessibility bias. The larger age difference in the high-interference condition does not reflect a deficit in ability to inhibit a prepotent response but, rather, arises from age differences in recollection having more opportunity to show themselves under conditions of high interference. Age differences in ability to inhibit are likely to exist, but to show those differences, effects of accessibility bias must be taken into account (Jacoby, Bishara, et al., 2005).

The low-interference condition in the above example, with accessibility bias set at .8, would be better described as a facilitation condition. Although we have focused on interference effects in the current article, the distinction between recollection and accessibility bias is equally important for measuring facilitation effects (learning). In our experiments, the retroactive facilitation gained by the List 1 target response being the most common response in the interpolated list was because of an advantage in accessibility bias, with recollection being unchanged. The issues for measuring facilitation effects are the same as those for measuring interference effects.

We began this article by asking the reader to recall the grade that was earned in his or her first high school math course and noting that memory distortion in answers to such questions can arise from proactive or retroactive probabilistic interference. Effects of that sort have generally been interpreted as reflecting the influence of schemas or implicit theories on the construction of the past (e.g., Bahrick et al., 1996; Ross, 1989). Claims that the past is constructed might imply that recall is accomplished by a generate/recognize process. However, as shown by our results, a viable alternative is that schemas and implicit theories can reflect probabilistic interference that contributes to accessibility bias, leaving ability to recollect unchanged.

Concluding Comments

Traditional investigations of learning and interference effects have used extreme probabilities of overlap in the pairing of cues and responses between List 1 and List 2. Perfect overlap between the two lists (A-B, A-B) constitutes an investigation of learning, whereas zero overlap between the two lists (A-B, A-D) constitutes an investigation of interference effects. By using intermediate probabilities of overlap, we have highlighted the possibility that response bias effects contribute to learning and interference. We have argued that it is necessary to distinguish between recollection and accessibility bias to understand effects of learning and interference on cued-recall performance just as it is necessary to distinguish between response bias and differences in discriminability to understand effects of response probability manipulations on recognition-memory performance.

Although there is controversy regarding the means by which discriminability and bias effects should be measured for recognition-memory performance (e.g., Snodgrass & Corwin, 1988), there is general agreement that it is necessary to distinguish between the two types of effect. As an example, one would draw very different conclusions about age-related differences in memory if older adults differed from young adults only in their response bias rather than, as typically found, differing in both response bias and discriminability (e.g., Craik & Jennings, 1992). In a related vein, we have shown that RI and PI sometimes only occur because of an influence on accessibility bias, with recollection being unchanged, and doing so impacts on the conclusions that can be drawn about RI and PI.

The assumptions underlying our estimation procedures are likely to hold only with some materials and across a limited set of conditions. An indirect test does not always serve as a process-pure measure of accessibility bias (e.g., Jacoby, 1991), and it is unlikely that recollection and accessibility bias are always fully independent bases for responding. Even if the underlying assumptions do not hold across a wide range of situations, having a situation in which they do hold is useful for purposes such as analyzing age-related differences in memory performance. A situation in which the assumptions hold also serves as a starting point for specifying the conditions under which they do not hold, which is important for theorizing about dissociations between indirect and direct tests of memory as well as for theorizing about interference effects.

There is strong support for the assumptions underlying the estimation procedures used in the current experiments. Results from our experiments indicate that sometimes, an indirect test can serve as a process-pure measure of accessibility bias (implicit memory). In contrast, our results suggest that performance on a direct test of memory is seldom, if ever, process pure. Rather, the recollection measured by direct test performance is contaminated by implicit memory in the form of accessibility bias. Separating contributions of different types of processes, as done with the recollection/accessibility bias model, holds advantages over trying to devise a direct test that provides a process-pure measure of recollection, which might be an impossible task.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19,* 716–723.

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

Anderson, M. C., & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Memory* (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.

Ayers, M. S., & Reder, L. M. (1998). A theoretical review of the misinformation effect: Predictions from an activation-based memory model. *Psychonomic Bulletin and Review, 5,* 1–21.

Bahrick, H. P., Hall, L. K., & Berger, S. A. (1996). Accuracy and distortion in memory for high school grades. *Psychological Science, 7,* 265–271.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge, England: Cambridge University Press.

Bodner, G. E., Masson, M. E. J., & Caldwell, J. I. (2000). Evidence for a generate-recognize model of episodic influences on word-stem completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 267–293.

Brewer, W. F. (2000). Bartlett's concept of schema and its impact on theories of knowledge representation in contemporary cognitive psychology. In A. Saito (Ed.), *Bartlett, culture and cognition* (pp. 69–89). New York: Psychology Press.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods and Research, 32,* 261–304.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (rev. ed.). New York: Academic Press.

Craik, F. I. M., & Jennings, J. M. (1992). Human memory. In F. I. M. Craik

& T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 51–110). Hillsdale, NJ: Erlbaum.

Curran, T., & Hintzman, D. L. (1997). Consequences and causes of correlations in process dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 496–504.

Dodson, C. S., Prinzmetal, W., & Shimamura, A. P. (1998). Using Excel to estimate parameters from observed data: An example from source memory data. *Behavior Research Methods, Instruments and Computers, 30,* 517–526.

Faul, F., & Erdfelder, E. (1992). GPOWER: A priori, post-hoc, and compromise power analyses for MS-DOS [Computer program]. Bonn, Germany: Bonn University, Department of Psychology.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review of a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 193–225). New York: Academic Press.

Hay, J. F., & Jacoby, L. L. (1999). Separating habit and recollection in young and elderly adults: Effect of elaborative processing and distinctiveness. *Psychology and Aging, 14,* 122–134.

Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychological Review, 96,* 208–233.

Humphreys, M. S., Dennis, S., Maguire, A. M., Reynolds, K., Boland, S. W., & Hughes, J. D. (2003). What you get out of memory depends on the question you ask. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29,* 797–812.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513–541.

Jacoby, L. L. (1996). Dissociating automatic and consciously-controlled effects of study/test compatibility. *Journal of Memory and Language, 35,* 32–52.

Jacoby, L. L., Begg, I. M., & Toth, J. P. (1997). In defense of functional independence: Violations of assumptions underlying the process-dissociation procedure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 484–495.

Jacoby, L. L., Bishara, A. J., Hessels, S., & Toth, J. P. (2005). Aging, subjective experience, and cognitive control: Dramatic false remembering by older adults. *Journal of Experimental Psychology: General, 134,* 131–148.

Jacoby, L. L., Debner, J. A., & Hay, J. F. (2001). Proactive interference, accessibility bias, and process dissociations: Valid subjective reports of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 686–700.

Jacoby, L. L., & Hollingshead, A. (1990). Toward a generate/recognize model of performance on direct and indirect tests of memory. *Journal of Memory and Language, 29,* 433–454.

Jacoby, L. L., Shimizu, Y., Velanova, K., & Rhodes, M. G. (2005). Age differences in depth of retrieval: Memory for foils. *Journal of Memory and Language, 52,* 493–504.

Jacoby, L. L., & Shrout, P. E. (1997). Toward a psychometric analysis of violations of the independence assumption in process dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 505–510.

Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General, 122,* 139–154.

Kelley, C. M., & Lindsay, D. S. (1996). Conscious and unconscious forms of memory. In E. L. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 31–63). San Diego, CA: Academic Press.

Kuha, J. (2004). AIC and BIC: Comparisons of assumptions and performance. *Sociological Methods and Research, 33,* 188–229.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13,* 585–589.

Luce, R. D. (1959). *Individual choice behavior.* New York: Wiley.

McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology, 114,* 1–16.

McFarland, C., & Ross, M. (1987). The relation between current impressions and memories of self and dating partners. *Personality and Social Psychology Bulletin, 13,* 228–238.

McGeoch, J. A. (1929). The influence of degree of learning upon retroactive inhibition. *American Journal of Psychology, 41,* 252–262.

Melton, A. W., & Irwin, J. M. (1940). The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *American Journal of Psychology, 53,* 173–203.

Mensink, G. M., & Raaijmakers, J. G. (1989). A model for interference and forgetting. *Psychological Review, 95,* 434–455.

Nelson, D. L., McKinney, V. M., Gee, N. R., & Janczura, G. A. (1998). Interpreting the influence of implicitly activated memories on recognition and recall. *Psychological Review, 105,* 299–324.

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

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

Reingold, E. M., & Toth, J. P. (1996). Process dissociations versus task dissociations: A controversy in progress. In G. D. M. Underwood (Ed.), *Implicit cognition* (pp. 159–202). New York: Oxford University Press.

Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology, 39,* 475–543.

Roediger, H. L., III, & McDermott, K. B. (1993). Implicit memory in normal human subjects. In H. Spinnler & F. Boller (Eds.), *Handbook of neuropsychology* (Vol. 8, pp. 63–131). Amsterdam: Elsevier.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review, 96,* 341–357.

Schneider, W. (1990). Micro experimental laboratory (Version 1.0) [Computer program]. Pittsburgh, PA: Psychology Software Tools.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6,* 461–464.

Shimamura, A. P. (1995). Memory and the prefrontal cortex. In J. Grafman, K. J. Holyoak, & F. Boller (Eds.), *Structure and functions of the human prefrontal cortex* (pp. 151–159). New York: New York Academy of Sciences.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117,* 34–50.

Warrington, E. K., & Weiskrantz, L. (1970, November 14). Amnesic syndrome: Consolidation or retrieval? *Nature, 228,* 628–630.

Winograd, E. (1968). List differentiation as a function of frequency and retention interval. *Journal of Experimental Psychology, 76,* 1–18.

Received November 8, 2005
Revision received November 15, 2006
Accepted November 16, 2006 ■