

January 1995

A comparison of item and source forgetting

Brian H. Bornstein

University of Nebraska-Lincoln, bbornstein2@unl.edu

Denny C. LeCompte

Louisiana State University, Baton Rouge, Louisiana

Follow this and additional works at: <http://digitalcommons.unl.edu/psychfacpub>



Part of the [Psychiatry and Psychology Commons](#)

Bornstein, Brian H. and LeCompte, Denny C., "A comparison of item and source forgetting" (1995). *Faculty Publications, Department of Psychology*. 327.

<http://digitalcommons.unl.edu/psychfacpub/327>

This Article is brought to you for free and open access by the Psychology, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications, Department of Psychology by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

A comparison of item and source forgetting

BRIAN H. BORNSTEIN and DENNY C. LECOMPTE
Louisiana State University, Baton Rouge, Louisiana

The purpose of the present research was to compare memory for an item with memory for the item's source. Experiment 1 investigated discrimination between two external sources: each item in a list of words was spoken in either a male or a female voice. Subjects received a test of item recognition and a test of source monitoring at each of four delay intervals (immediate, 30 min, 48 h, 1 week). In contrast with previous research, no evidence of differential forgetting rates for item and source information was found. With delay intervals of 0 and 48 h, Experiment 2 replicated Experiment 1 while adding a reality monitoring condition that required discrimination between an internal (i.e., self-generated) and an external source. Subjects were better at making internal-external discriminations than at making external-external discriminations, but both types of source monitoring declined at the same rate as memory for the items themselves.

A dissociation can be made between memory for an *item* and memory for its origin, or *source* (Johnson, Hashtroudi, & Lindsay, 1993). For example, one might correctly remember that the African National Congress won the first open South African election without remembering if the information was acquired from television, radio, newspaper, or a friend. Memory for source information—often referred to as “source monitoring” (see, e.g., Johnson et al., 1993)—is important in a variety of situations and domains. This is most clearly demonstrated by instances in which source memory fails, as in telling a joke thought to be one's own invention to the person from whom one first heard it, or confusing the memory of a crime event with information about the crime that became available subsequently, such as whether or not a weapon was present.

Studies collecting data on both source and item memory have yielded a consistent finding: item information is remembered better than source information (see, e.g., Batchelder & Riefer, 1990; Brown & Halliday, 1991; Voss, Vonder, Post, & Ney, 1987). The inferiority of source memory is especially pronounced in older adults (e.g., Cohen & Faulkner, 1989; Hashtroudi, Johnson, & Chrosniak, 1989; McIntyre & Craik, 1987) and patients with organic memory disorders (Schacter, Harbluk, & McLachlan, 1984; Shimamura & Squire, 1987), suggesting that the ability to remember items and the ability to remember the source of those items are dissociable processes.

Although it is well established that item information is remembered better than source information overall, the

rate at which the two types of information are forgotten is less clear. Studies of source monitoring that investigate the rate of forgetting vary widely in terms of the experimental paradigm and delay interval used. Not surprisingly, source information is remembered less well 1 week after presentation than after only 10 min, in both young and older adults (McIntyre & Craik, 1987, Experiment 2). Furthermore, Craik and Kirsner (1974, Experiments 3 and 4) and Marsh and Bower (1993, Experiment 1) both found that source-monitoring performance deteriorates significantly within just a few minutes of presentation. However, neither Marsh and Bower nor Craik and his colleagues made direct comparisons between source and item memory, or attempted to draw any conclusions about the relative effect of delay on both processes.

In contrast, Brown and Halliday (1991) directly compared source and item forgetting, immediately and after a 1-week delay. Using the paradigm developed by Brown and Murphy (1989), they had members of a group take turns generating category exemplars (e.g., different kinds of fruit); later, each subject classified the exemplars on a recognition test (either immediately or 1 week later) as their own, as others', or as new responses. Brown and Halliday were thus able to measure instances of both source forgetting—defined as “own” responses called “others” and “others” responses called “own”—and item forgetting, defined as “own” and “others” responses called “new.” Based on these indices, both types of forgetting increased significantly over time, but source forgetting increased at a faster rate (the ratio of source to item errors was 2.3:1 at immediate test and 3.2:1 at 1-week delay). Thus, they conclude that source information is forgotten at a faster rate than is item information.

This conclusion is limited by two factors, both of which derive from Brown and Halliday's (1991) choice of paradigm. First, the same recognition test was used to measure both item and source memory, meaning that source memory can be assessed only if an item is identified as “old.” Using this technique, source memory may be mea-

We are grateful to Angelle Bourgeois, Craig Neely, Joel Ringdahl, Laurel Salley, and Jennifer Tabor for their assistance in creating the stimuli and gathering data, and to Alan Brown, John Gardiner, Stephanie LeCompte, and Janet McDonald for helpful comments on the manuscript. The authors contributed equally to all phases of the project. Correspondence may be addressed to either author at 236 Audubon Hall, Department of Psychology, Louisiana State University, Baton Rouge, LA 70803 (e-mail: psbrian@lsuvm.sncc.lsu.edu or lecompte@rada.cogsci.lsu.edu).

sured for only a subset of items in the original presentation—that is, those correctly recognized as “old.” Second, Brown and Halliday (1991) included only “misses” in their item-forgetting index; by omitting false positives—“new” items called “own” or “others”—this index may underestimate the true incidence of item forgetting.

EXPERIMENT 1

Our first experiment was an attempt to replicate Brown and Halliday’s (1991) findings with a task requiring subjects to distinguish between two external sources as opposed to making an internal–external discrimination. We solved the problem of using the same test to measure two dissociable processes by using different stimuli for separate item and source tests in a within-subject design. The within-subject method allows for a more sensitive comparison between item and source memory, especially regarding their relative rate of forgetting. The delay between stimulus presentation and test was also manipulated within subjects, using more levels of delay than had been used in any previous study, thereby allowing for a finer assessment of the forgetting function.

Specifically, subjects were presented a list of words, each of which was read in either a male or a female voice. The subjects’ memory for the list items and their memory for the source of the items were tested immediately and at delays of 30 min, 48 h, and 1 week. We made three predictions: (1) Subjects’ ability to recognize that an item had been on the list would exceed their ability to recognize its source—that is, which voice had spoken the item. As noted earlier, this is a common finding (e.g., Brown & Halliday, 1991; Voss et al., 1987) and may derive from the fact that source monitoring requires, on average, more differentiated input (Johnson et al., 1993). (2) Both item recognition and source monitoring would decrease as the delay increased. Like other aspects of memory, source monitoring should deteriorate over time (cf. Marsh & Bower, 1993). (3) The rate of forgetting would be faster for source information than for item information. We sought to extend Brown and Halliday’s (1991) finding to a different source-monitoring situation, with the additional procedural modifications of a within-subject design and a larger number of delay intervals.

Method

Subjects. The subjects were 48 Louisiana State University undergraduates.

Materials and Design. The stimuli were 160 words chosen randomly from the Toronto Word Pool (Friendly, Franklin, Hoffman, & Rubin, 1982). The words were divided randomly into two lists of 80 words, List A and List B. Each list of words was recorded onto a cassette tape to be played during the study phase of the experiment. Forty randomly chosen words were spoken in a male voice, and 40 were spoken in a female voice; the order of the male- and female-voiced words within each list was random, with the constraint that for each word spoken in a particular voice in List A, the word in the corresponding ordinal position in List B be spoken in the same voice. Half of the subjects heard List A; the other half heard List B.

The first test was given immediately after the study list; a second test was given 30 min later, a third was given after 48 h, and a fourth 168 h (1 week) after the study list.

Each word from the study list was tested once during the experiment. Each test phase consisted of 10 item-recognition test stimuli and 10 source-monitoring test stimuli. Each stimulus on the item test comprised a word from the study list paired with the word in the corresponding ordinal position from the unstudied list. For example, the first word on List A was *forgive* and the first word on List B was *player*; thus one of the item-test stimuli was the pair *forgive player*. Accordingly, each set of 10 item-test stimuli used 20 words from the total pool of 160 words, and the four item tests together used 80 words.

Each item on the source test comprised a single word from the study list followed by the letters M and F—for example, *forgive M F*. The four source tests used 10 words each for a total of 40 words from the initial pool of 160 words. Together, the item test and the source test accounted for 120 of the 160 stimulus words in the experiment. For each subject, the remaining 40 words were from the unstudied list that corresponded in position to the 40 words used for their source tests.

A number of counterbalancing measures were taken with the test lists. First, for half of the subjects, the first 10 items in each of the four tests were item stimuli and the second 10 were source stimuli. The order of item and source test stimuli was reversed for the other half of the subjects. Second, each word was used for an item test for half of the subjects and for a source test for the other half of the subjects. Third, the test stimuli were divided into four sets of 20 stimuli, 10 item and 10 source stimuli. These four sets were rotated through the four test times (0, 0.5, 48, and 168 h) according to a Latin square. This rotation ensured that test items were not confounded with test times. Fourth, and finally, the left/right position of each item-test stimulus was reversed for half of the subjects. All of these counterbalancing measures were combined to yield 16 separate counterbalancing conditions.

Procedure. For the study phase, the subjects were told that they would hear a series of words, some in a male voice, some in a female voice. The subjects were also told that they needed only to remember which voice spoke each word; they did not expect to be tested on their memory for the items themselves. To ensure that subjects attended to the study list, as each word was spoken, they indicated whether they could generate a mental image of it.¹

The list of 80 words was played over loudspeakers at a 5-sec rate. Immediately after the list was completed, the experimenter gave the instructions for both the item test and the source test. During this testing phase, each subject wore headphones attached to a small tape player that guided them through the test. The subjects had the written test in front of them while listening to the auditory version. For each item-test stimulus, the studied word and the unstudied word were spoken by the voice that had originally spoken the studied word, with a 2-sec pause separating the words. Test stimuli were spoken aloud because re-presenting a stimulus auditorily has been found to provide a more sensitive test of word recognition (Craig & Kirsner, 1974; Palmeri, Goldinger, & Pisoni, 1993). The words were always spoken in the order in which the items were presented on the page from left to right. The subjects circled the word from each pair that they thought was from the study list.

For each source-test stimulus, the word from the study list was spoken first in the male voice and then in the female voice, with a 2-sec pause separating the utterances. The subjects circled either M or F to indicate that they thought the word was originally spoken in either a male or female voice. Next, the subjects engaged in an unrelated short-term memory task until 30 min had elapsed since the end of the study list. The subjects were then given another item test and another source test.

Following the second pair of tests, the subjects were sent home. They returned 2 days later and again 5 days after that. During each of these sessions, the subjects were given an item test and a source test.

Results and Discussion

The results are shown in the left-hand panel of Figure 1. Item performance exceeded source performance [$F(1,47) = 5.32, p = .03$], confirming our first prediction.

Likewise, supporting our second prediction, performance declined from the first to the fourth test [$F(1,47) = 19.17, p < .001$] and planned linear contrasts showed that there was a significant decline in performance for both item recognition [$F(1,47) = 29.52, p < .001$] and source monitoring [$F(1,47) = 36.49, p < .001$].

Because of a problem with a floor effect in pilot testing (see Note 1), we compared both item recognition and source monitoring to what would be expected by chance (.5) at each of the four delay intervals. Performance reliably exceeded chance at every interval, both for item recognition and source monitoring [$t(47) > 1.9, p < .05$, in every case, by a one-tailed test]. Thus, the subjects exhibited the ability to remember both item and source information for as long as 1 week after initially studying the stimuli.

Our third prediction was that the rate of forgetting for source monitoring would exceed the rate of forgetting for item recognition. The simple interaction between type of test and delay failed to reach statistical significance [$F(3,141) = 0.35, p = .79$]; likewise, a comparison of the linear trends for item recognition and source monitoring revealed no discernible difference [$F(1,47) = 0.91, p = .34$], thereby failing to support the notion that memory for source information declines faster than item memory.

EXPERIMENT 2

Our finding that source information was not forgotten more rapidly than item information contradicts results reported previously by Brown and Halliday (1991). Our procedure deviated from theirs in that we used an external source-monitoring task rather than a reality monitoring task (Johnson et al., 1993). Reality monitoring involves determining whether information was self-generated or

had an external source (Johnson & Raye, 1981), as in Brown and Halliday's (1991) experimental task, where subjects attributed information either to themselves or to another subject. External source monitoring, on the other hand, requires discriminating among multiple external sources, as in Experiment 1, where subjects attributed information to either a male or a female speaker.²

Internally and externally derived memories differ along a number of dimensions, such as the amount of sensory information and information about cognitive operations (Johnson et al., 1993; Johnson & Raye, 1981). These differences lead to the prediction that discriminating between items belonging to a single class (i.e., internal-internal or external-external) will be more difficult than discriminating between classes (Johnson et al., 1993). This prediction has been supported by Raye and Johnson (1980), who found that subjects were better at discriminating between an internal and an external source than between two external sources, an advantage that becomes even greater in elderly subjects (Hashtroudi et al., 1989).

Although reality monitoring is superior to external source monitoring overall, the relative rate of forgetting for these two types of source monitoring has not yet been investigated. The same stimulus characteristics which make reality monitoring easier than external source monitoring might also influence forgetting in the two types of tasks, thereby accounting for the discrepancy between the results of Experiment 1 and those obtained by Brown and Halliday (1991). The ability to distinguish between sources might be lost more quickly when the distinction is between an internal and an external source (i.e., reality monitoring) than when it is between two external sources. Experiment 2 tested the possibility that reality monitoring would show a faster rate of forgetting than item memory. It also included an external-source moni-

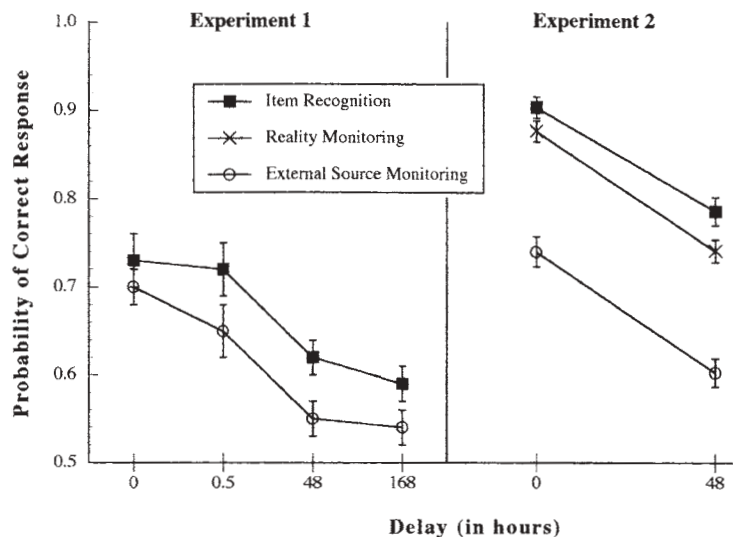


Figure 1. Probability of item recognition and [external] source monitoring in Experiment 1 (left panel) and of item recognition, external source monitoring, and reality monitoring in Experiment 2 (right panel). Error bars represent standard errors.

toring task, to allow for a replication of Experiment 1 and to provide a direct comparison of the two types of source-monitoring tasks.

As in Experiment 1, the experiment consisted of both a study phase and a test phase. The study phase comprised a list of words, each of which was presented in one of three different ways: auditorily, visually, or as a word fragment requiring subjects to generate the word. The test phase was similar to that of Experiment 1 in that subjects made one of several different judgments for each word on the study list. As in Experiment 1, for some of the words, subjects were to make an item-recognition judgment—that is, which of two words (target or distractor) had been presented during the study phase; for other words, they were to make an external source-monitoring judgment—that is, was the word presented visually or auditorily. In addition, for some of the words, the subjects were to make a reality-monitoring judgment—that is, was the word generated as a word fragment or presented externally (either visually or auditorily).

Method

Subjects. The subjects were 66 Louisiana State University undergraduates who received extra course credit in exchange for their participation.

Materials and Design. The experiment was a 2×3 within-subject design. There were two test delays, immediate and 48 h, and three types of test, item recognition (old vs. new), external source monitoring (auditory vs. visual), and reality monitoring (internally vs. externally generated).

From each of 24 categories in the Battig and Montague (1969) stimulus set (e.g., *color*), we selected the eight most typical exemplars, exempting those that comprised two words or that belonged to more than 1 category. For each category, the eight exemplars were divided into four pairs consisting of the first and second, third and fourth, fifth and sixth, and seventh and eighth most typical exemplars, respectively. For each pair, one exemplar was randomly assigned to serve as a target and one to serve as the distractor for that target in the item-recognition test (described below). Thus, there were 96 targets and 96 distractors.

Study phase. Each subject sat at an Apple Macintosh Classic II computer and wore headphones attached to the computer. As in Experiment 1, the subjects were instructed to attend only to source information.

Each word was presented in one of three different ways: auditorily, visually, or as a word fragment requiring subjects to generate the response. For auditory presentation, each word was digitally recorded in a male voice and presented through headphones. For visual presentation, each word was presented in 24-point Geneva font. For presentation as word fragments, a word fragment was created for each target by removing all of the even-numbered letters (e.g., *blue* became *b_u_*); however, for exemplars longer than four letters, we retained the second letter (e.g., *yellow* became *yel_o_*), because pilot testing had suggested that many exemplars become too difficult otherwise. Likewise, extra letters were retained in a handful of relatively long exemplars. The 96 target words were assigned randomly and evenly to the three presentation modes, and this randomization was performed separately for each subject.

To constrain the responses to the word fragments, a category name (e.g., *colors*) appeared on the computer screen 0.5 sec before the word fragment. For the sake of consistency, a category name preceded auditory and visual presentations as well. When subjects were presented with a word fragment, they typed their solution to it. If their solution was correct, the study list continued. If it was incorrect, they were allowed to try another solution. At this point, if there was more than one missing letter, a randomly chosen letter was replaced. After three unsuccessful attempts, the subjects saw a message instructing them to ask the experimenter for a hint. This happened infrequently (less than once

per subject), and in each case, the experimenter gave the subject a cue which allowed him or her to complete the fragment (e.g., “rhymes with ‘glue’”). In every case, the subject was eventually able to complete the fragment.

When the subjects saw or heard an exemplar, they had to wait at least 3 sec before the computer would accept a keypress that would continue the study list. Because the study phase was self-paced, the subjects spent more time and effort, on average, on generating the word fragments than on hearing or seeing the other two types of words. This confound is inherent to research using a reality-monitoring paradigm (Johnson & Raye, 1981) and is minimized by Slamecka and Graf’s (1978) finding that the size of the generation effect is roughly the same for self-paced and experimentally controlled presentation.

Test phase. When the subjects completed the study list, they were informed about the test phase, which consisted of a series of items presented one at a time on the computer screen. Each test item consisted of one of three forced-choice judgments about a word that had been presented during the study phase: item recognition, reality monitoring, or external source monitoring. For the item-recognition test, the subjects saw a target and a distractor and selected the item they thought had been presented in the study phase. For the reality-monitoring test, the subjects saw a target and selected either “typed by you” (i.e., generated) or “saw/heard.” For the external source-monitoring test, they saw a target that had been presented externally and selected either “saw” or “heard.”

For each judgment, one alternative was presented on the left of the screen and the other was presented on the right of the screen. For each subject, the left–right order was determined randomly and separately for each item. The subjects had to choose one of the two alternatives before they could go on to the next test item. The three test types were mixed together randomly and separately for each subject. The item-recognition test contained 4 visual, 4 auditory, and 8 generated targets (plus 16 distractors); the reality-monitoring test contained 4 visual, 4 auditory, and 8 generated targets; and the external source-monitoring test contained 8 auditory and 8 visual targets.

After responding to all of the test items, the subjects were sent home. They returned 48 h later and took the second test, which was procedurally identical to the first. Within each test type, half of the target items were tested immediately and half were tested after 48 h. Thus, there were 48 test items at each delay, 16 for each type of test. The assignment of each target to immediate or 48-h testing was determined randomly for each subject. Furthermore, within the immediate and 48-h testing conditions, the order of the exemplars was also random and separate for each subject.

Results and Discussion

The results are illustrated in the right-hand panel of Figure 1.³ There was a main effect of type of test [$F(2,130) = 84.08, p < .001$]. As in Experiment 1, item-recognition performance reliably exceeded external source-monitoring performance [$F(1,130) = 151.17, p < .001$], and, replicating Raye and Johnson (1980), reality-monitoring performance reliably exceeded external source-monitoring performance as well [$F(1,130) = 94.43, p < .001$]. Finally, as found by Brown and Halliday (1991), item-recognition performance reliably exceeded reality-monitoring performance [$F(1,130) = 6.65, p = .01$].

There was also a main effect of delay [$F(1,65) = 155.31, p < .001$], and planned comparisons revealed a statistically significant decline in performance for the item-recognition test, the reality-monitoring test, and the external source-monitoring test [$F(1,65) = 42.88, 56.92, \text{ and } 56.92, \text{ respectively; } p < .001 \text{ in each case}$].

Of primary concern, however, is the interaction between type of test and delay. The overall interaction was not statistically significant [$F(2,130) = 0.33, p = .72$].

Planned comparisons testing for an interaction between delay and each pair of tests were also conducted. As in Experiment 1, the effect of delay did not interact with type of test for item recognition and external source monitoring [$F(1,130) = 0.50, p = .48$]. The reality-monitoring test also failed to show a steeper rate of forgetting than the item-recognition test [$F(1,130) = 0.50, p = .48$]. Finally, the difference in rate of forgetting between the external source-monitoring and reality-monitoring conditions was also nonsignificant [$F(1,130) < 0.01$]. Thus, there is no evidence that source information is forgotten faster than item information, regardless of whether the source information involves a discrimination between two external sources or between an internal and an external source.

GENERAL DISCUSSION

As previous research has demonstrated (e.g., Brown & Halliday, 1991; Voss et al., 1987), subjects recognize whether or not an item was previously encountered in an experimental situation better than they recognize the source of that item. Item recognition is superior whether the source monitoring involves discrimination between an internal and an external source or between two external sources. As suggested by Johnson et al. (1993), any type of source monitoring might require, on average, more differentiated input than item recognition. This general requirement would produce a tendency toward inferior source-monitoring performance. Even when subjects were instructed to attend only to source information and the original voice was re-presented during test in order to improve their source monitoring (Experiment 1), their item recognition was still consistently superior. Of the two types of source monitoring, subjects performed better on the reality-monitoring task (Experiment 2), supporting Johnson and Raye's (1981; Raye & Johnson, 1980) theory regarding the general differences between internally and externally derived memories.

One of the most robust findings in all memory research is that memory diminishes over time, and source memory is no different in this respect. Our results are consistent with previous findings (Brown & Halliday, 1991; Craik & Kirsner, 1974; Marsh & Bower, 1993) showing that source monitoring, like other aspects of memory, deteriorates over time. If source memory begins at a sufficiently high level, incremental forgetting of source information continues for periods as long as 1 week; nevertheless, after 1 week, source monitoring is still better than chance.

In contrast to previous findings (Brown & Halliday, 1991), we obtained no evidence that memory for an item's source—in terms of either reality monitoring or external source monitoring—declines more rapidly than memory for the item itself. Our method differs from Brown and Halliday's in that we used a within-subject, rather than a between-subjects, design. The within-subject design was chosen in order to use different stimuli for separate item and source tests, thereby overcoming the problem of using the same test to measure both processes (i.e., "given that the information is correctly recognized as 'old,' what is its source?"). In using the same test to measure both item and source forgetting, Brown and Halliday did not consider false positives (i.e., "new" items called "own" or "others") to be evidence of item forgetting, even though these errors show just as much inability to distinguish between old and new items as does mislabeling an old item as "new." Thus, the true incidence of item forgetting is potentially higher than their index indicates. Using this modified testing procedure, we obtained equivalent forgetting rates for item recognition, external source monitoring, and reality monitoring.⁴

In summary, we replicated previous research in finding that item information is remembered better than source information, whether the source monitoring involves discriminating between two external sources or between an internal and an external source. Of the two types

of source monitoring, subjects are better at making internal-external distinctions. Also consistent with previous research, performance on all three tasks declines over time. However, we obtained no evidence that memory for an item's source declines more rapidly than memory for the item itself.

REFERENCES

- BATCHELDER, W. H., & RIEFER, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review*, *97*, 548-564.
- BATTIG, W. F., & MONTAGUE, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monographs*, *80*(3, Pt. 2).
- BROWN, A. S., & HALLIDAY, H. E. (1991). Cryptomnesia and source memory difficulties. *American Journal of Psychology*, *104*, 475-490.
- BROWN, A. S., & MURPHY, D. R. (1989). Cryptomnesia: Delineating inadvertent plagiarism. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *15*, 432-442.
- COHEN, G., & FAULKNER, D. (1989). Age differences in source forgetting: Effects on reality monitoring and on eyewitness testimony. *Psychology & Aging*, *4*, 10-17.
- CRAIK, F. I. M., & KIRSNER, K. (1974). The effect of speaker's voice on word recognition. *Quarterly Journal of Experimental Psychology*, *26*, 274-284.
- FRIENDLY, M., FRANKLIN, P. E., HOFFMAN, D., & RUBIN, D. C. (1982). The Toronto Word Pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods & Instrumentation*, *14*, 375-399.
- HASHTRUDI, S., JOHNSON, M. K., & CHROSNIAK, L. D. (1989). Aging and source monitoring. *Psychology & Aging*, *4*, 106-112.
- HYDE, T. S., & JENKINS, J. J. (1969). Differential effects of incidental tasks on the organization of recall of a list of highly associated words. *Journal of Experimental Psychology*, *83*, 472-481.
- JOHNSON, M. K., HASHTRUDI, S., & LINDSAY, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3-28.
- JOHNSON, M. K., & RAYE, C. L. (1981). Reality monitoring. *Psychological Review*, *88*, 67-85.
- LOFTUS, G. R. (1985). Evaluating forgetting curves. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *11*, 397-406.
- MARSH, R. L., & BOWER, G. H. (1993). Eliciting cryptomnesia: Unconscious plagiarism in a puzzle task. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *19*, 673-688.
- MCINTYRE, J. S., & CRAIK, F. I. M. (1987). Age differences in memory for item and source information. *Canadian Journal of Psychology*, *41*, 175-192.
- PALMERI, T. J., GOLDINGER, S. D., & PISONI, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *19*, 309-328.
- RAYE, C. L., & JOHNSON, M. K. (1980). Reality monitoring vs. discriminating between external sources of memories. *Bulletin of the Psychonomic Society*, *15*, 405-408.
- SCHACTER, D. L., HARBLUK, J. L., & MCLACHLAN, D. R. (1984). Retrieval without recollection: An experimental analysis of source amnesia. *Journal of Verbal Learning & Verbal Behavior*, *23*, 593-611.
- SHIMAMURA, A. P., & SQUIRE, L. R. (1987). A neuropsychological study of fact memory and source amnesia. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *13*, 464-473.
- SLAMECKA, N. J. (1985). On comparing rates of forgetting: Comment on Loftus (1985). *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *11*, 812-816.
- SLAMECKA, N. J., & GRAF, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning & Memory*, *4*, 592-604.
- VOSS, J. F., VESONDER, G. T., POST, T. A., & NEY, L. G. (1987). Was the item recalled and if so by whom? *Journal of Memory & Language*, *26*, 466-479.

WIXTED, J. T. (1990). Analyzing the empirical course of forgetting. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *16*, 927-935.

NOTES

1. These instructions were chosen on the basis of pilot testing. If subjects were instructed to remember both the words and the voice in which each word was spoken, their performance on the source test did not differ from chance after just a 30-min delay. Although the instructions confound the source-item comparison with intentionality, research has shown that incidental learning is roughly equivalent to intentional learning given a relatively "deep" orienting task (e.g., Hyde & Jenkins, 1969), such as the imagery instruction used in Experiment 1.

2. This distinction between reality monitoring and external source monitoring roughly follows the framework proposed by Johnson et al. (1993), who go on to distinguish two different forms of reality monitoring: (1) reality monitoring, per se, which compares internal (performed or imagined) and external (perceived) events, and (2) discriminating between what one merely thought (imagined) and what one

actually said or did (performed), labeled "internal source monitoring." The two subtypes are alike, and distinguished from external source monitoring, in their inclusion of the self as a potential source of information. For the purposes of this paper, "reality monitoring" refers to the former type—that is, making an internal-external distinction.

3. Subjects appear to have done considerably better on the item-recognition test in Experiment 2 than in Experiment 1. This can be attributed to the inclusion in Experiment 2 of the word-fragment condition, which would produce a generation effect in remembering which stimuli had been on the study list (Slamecka & Graf, 1978).

4. In comparing forgetting rates, we have employed what Slamecka (1985) has referred to as the "customary" procedure. It must be acknowledged that there are other means of measuring forgetting (e.g., Loftus, 1985); however, these alternative methods require the a priori adoption of a particular model of forgetting. Following Slamecka (1985; see also Wixted, 1990), we have taken a primarily functional approach to the issue of forgetting rates.

(Manuscript received July 5, 1994;
revision accepted for publication December 22, 1994.)