Journal of Experimental Psychology: General
1996, Vol. 125, No. 2, 181–194

0096-3445/96/$3.00

Some Problems With the Process-Dissociation Approach to Memory

Chad S. Dodson and Marcia K. Johnson
Princeton University

The process-dissociation framework (L. L. Jacoby, 1991) is a technique for deriving estimates of controlled (e.g., recollection) and automatic (e.g., familiarity) memory processes. The authors examined 3 assumptions of this framework. In Experiment 1, estimates of familiarity were affected by varying the proportion of old targets to old nontargets on the inclusion and exclusion tests and whether or not the tests were completed with full or divided attention, violating the assumption that familiarity's influence is automatic. In Experiment 2, the similarity of old targets and old nontargets was manipulated to show that source confusions (i.e., misrecollections) violate the assumption that the process of recollection is all-or-none. Source confusions also create an imbalance in the influence of recollection on the inclusion and exclusion tests, violating the consistency assumption. The source-monitoring framework is consistent with the present findings.

Theories of recognition generally posit two factors. The first is typically characterized as a feeling of familiarity that is based on some unidimensional variable resulting from the frequency of prior exposure (e.g., Atkinson & Juola, 1974; Underwood, 1972), the fluency of processing the recognized item (e.g., Jacoby & Dallas, 1981), the integration of the underlying memory representation (e.g., Mandler, 1980), or the similarity of the item to memory (e.g., Gillund & Shiffrin, 1984; Hintzman, 1988; Humphreys, Bain, & Pike, 1989; Murdock, 1982). The second, often referred to as retrieval or recollection, is described as involving more information, such as multiple attributes of an event; an entire remembered event and the context in which it occurred; or an event and associated, elaborative information (e.g., Anderson & Bower, 1972, 1974; Gillund & Shiffrin, 1984; Humphreys, et al., 1989; Mandler, 1980).

Most theorists have agreed that familiarity emerges relatively more automatically (typically is faster, takes less effort, is less susceptible to disruption from competing tasks, etc.) than does the recollection of more specific information (e.g., Atkinson & Juola, 1974; Hintzman & Curran, 1994; Mandler, 1980). Jacoby and colleagues (e.g., Jacoby, 1991; Jacoby & Kelley, 1992; Jacoby, Lindsay, & Toth, 1992; Jacoby, Toth, & Yonelinas, 1993) went a step farther and equated familiarity with automaticity and recollection with control (e.g., Hasher & Zacks, 1979; Posner & Snyder, 1975; Schneider & Shiffrin, 1977). They have proposed a technique—*the process-dissociation framework*—for deriving quantitative estimates of the contributions of automatic and controlled processes to a given task. The basic idea of the process-dissociation framework is that familiarity arises and is used automatically, whereas recollection of a particular event arises and is used in a controlled manner. Furthermore, it is assumed that familiarity may lead to correct or incorrect recognition judgments, depending on task conditions, but that recollection always leads to a correct response. Thus the inappropriate influence of familiarity can be opposed by recollection to eliminate incorrect responses.

An alternative view of the roles of familiarity and recollection in remembering derives from the source-monitoring framework proposed by Johnson and colleagues (Johnson, Hashtroudi, & Lindsay, 1993; Johnson & Hirst, 1993). The source-monitoring framework assumes that there are varieties of memorial information (e.g., semantic, perceptual, affective) that are accessed either directly by cues or with the help of reflective processes; which information is activated, and for how long, depends not only on available cues but also on an individual's goals. This activated information gives rise to various types of phenomenal experiences ranging in specificity from a vague feeling of familiarity to a vivid recollection of an event. This framework distinguishes familiarity from recollection in terms of the specificity of the memorial information (i.e., how vague or specific it is) and not in terms of automatic and controlled processes. Remembering involves decision-making (or attributional) processes that flexibly make use of the amount and kind of memorial information that is available depending on task demands and the goals of the rememberer (e.g., Dodson & Johnson, 1993; Johnson et al., 1993; Lindsay & Johnson, 1989; Raye, 1976). These decision-making processes sometimes fail or sometimes operate on memorial information that is not specific enough to identify the source of an item, resulting in various kinds of source confusions or source misattributions.

Chad S. Dodson and Marcia K. Johnson, Department of Psychology, Princeton University.

This research was funded in part by a National Science Foundation Graduate Fellowship and National Institutes on Aging Grant AG09253. For their helpful comments, we thank Sam Glucksberg, Deborah Prentice, Ron Kinchla, Lawrie Balfour, Mike Anderson, Ute Bayen and, especially, the past and present members of the Princeton Memory Lab: Jennifer Bonilla, Barb Chalfonte, Allison Hermann, Tina Loose, Kristi Multhaup, Scott Nolde, John Reeder, Sarah Smith, and Christine Voegeli.

Correspondence concerning this article should be addressed to Chad S. Dodson, who is now at the Department of Psychology, University of California, Berkeley, California 94720. Electronic mail may be sent via Internet to chadd@violet.berkeley.edu.

For the present discussion, an important feature of the source-monitoring approach is that the kind of information that is activated is orthogonal to whether or not its influence is automatic or controlled. Both familiarity and more specific information may arise from and be used in either more automatic or more controlled (reflective) processing (or both; e.g., Johnson & Hirst, 1993). With respect to the activation of information, we agree that the experience of familiarity often arises from encounters with stimuli without additional reflective effort (i.e., automatically) and that remembering source-specifying information often requires reflectively guided retrieval (i.e., controlled) processes. However, the opposite occurs as well. Highly specific memories may be directly elicited by cues, as in the sudden memory of a dream prompted by an appropriate cue. Conversely, effortful search may result only in a feeling of greater familiarity without any increase in specific information, as when one dwells on the answers to a difficult multiple-choice question. Thus both familiarity and more specific memory can arise relatively automatically or after some effort.

Familiarity and more specific information not only can arise through automatic or controlled processes but also may be used through heuristic (relatively automatic) or strategic (relatively controlled) processes for some particular task or agenda, such as making old–new discriminations or identifying items from various sources.[1] For example, consider a person at a conference who seems familiar but about whom you cannot remember any contextual information, such as a name, voice, or research area. You may feel that the probability is high that you have met the person before and, hence, be reluctant to act as if you do not recognize them (e.g., by introducing yourself and asking their name). Therefore, you act as if you recognize the person. On the other hand, if you have the same familiarity response to the same person while on vacation in Spain, you may well introduce yourself and ask their name, because the benefits of friendly interaction in a strange place outweigh the risks of embarrassment at not recognizing someone you have met before (and the unusual context provides a ready excuse for not recognizing them). Thus, the threshold for an overt recognition response on the basis of familiarity alone is affected by the subjective probability of the target being old in combination with the agenda and other information that are active.

If familiarity is sometimes used strategically, as the above example and the source-monitoring framework suggest, then there are circumstances in which the process-dissociation framework is inappropriate. Consistent with this, we report results indicating three problems with the process-dissociation framework: First, familiarity's influence is not always automatic; second, the influences of familiarity and recollection are not invariant across situations in which invariance must be assumed in order to quantify the influences of familiarity and recollection; and third, the recollection of specific information is not an all-or-none process. Instead, there are varieties and degrees of recollected information that are more and less sufficient for particular tasks.

The Process-Dissociation Framework

The process-dissociation procedure consists of contrasting test situations in which recollection and familiarity work together to contribute to a memory judgment with test situations in which the two processes work in opposition. To illustrate, consider a study by Jacoby (1991) in which participants read words and solved anagrams in Phase 1 and then heard a list of words in Phase 2. Finally, some participants received an inclusion test in which all items that had been previously read, solved as an anagram, or heard were to be called *old,* and only the new items were to be called *new.* According to the process-dissociation framework, for each class of items (i.e., items read, presented as anagrams, and heard) on the inclusion test, both recollection (R) and familiarity (F) should jointly facilitate performance. Equation 1 summarizes the processing components that could contribute to the probability of saying a particular item was old on the inclusion test; a judgment of old can be based on recollecting the item, or in the absence of recollection, participants may base this judgment on the item's familiarity:

$$p(\text{correct recognition} \mid \text{inclusion test}) = \mathrm{R} + \mathrm{F}(1 - \mathrm{R}). \quad (1)$$

The remaining participants in Jacoby's study (1991) received an exclusion test in which participants were instructed to call old only the words that had been previously heard; all the other words that had been read, solved as an anagram, or that were new were to be called new. According to the process-dissociation framework, recollection and familiarity work in opposition in this task. If participants mistakenly called an item old that had actually been read or that had been presented as an anagram, then the mistake must have been due to the item's familiarity, unopposed by recollection. If participants had recollected that the item had been read or solved as an anagram, then they would have discounted its familiarity and correctly responded new. Equation 2 expresses the idea that familiarity, acting in the absence of recollection, accounts for the false recognition (i.e., incorrectly saying old) of a previously studied word that should have been rejected, such as an anagram:

$$p(\text{false recognition} \mid \text{exclusion test}) = \mathrm{F}(1 - \mathrm{R}). \quad (2)$$

As can be seen from examining Equations 1 and 2, the values of R and F for a particular type of item are easily determined from the correct-recognition rate on the inclusion task and the false-recognition rate on the exclusion task. Only the process of recollection can be used to select for items on the inclusion test and to select against these items on the exclusion test. The influence of familiarity, however, is not under participants' control. In the absence of recollection, familiarity should contribute to false alarms

[1] *Heuristic* and *strategic* are shorthand ways of referring to differences in the amount of reflective activity required by a task (Johnson et al., 1993). They are not opposites of a single dimension in the way that automatic and controlled processes are often assumed to be.

on the exclusion test to the same extent that it contributes to correct responses on the inclusion test.

Assumptions of the Process-Dissociation Framework

Jacoby (1991) has described three assumptions underlying the process-dissociation framework. First, the influence of familiarity on a particular class of items must be the same on the inclusion and exclusion tests. Similarly, recollection must contribute to performance to the same extent on both tests. We refer to these first two assumptions as the *consistency assumption.* Finally, it is assumed that recollection is completely independent from familiarity as a basis for a decision. The values of familiarity and recollection are assumed to be uncorrelated (however, see Curran & Hintzman, 1995; Joordens & Merikle, 1993).

The consistency assumption should be viewed with some suspicion because different memory tests can affect the kind and amount of memorial information that influence a memory judgment (e.g., Dodson & Johnson, 1993; Jacoby, Kelley, Brown, & Jasechko, 1989; Lindsay & Johnson, 1989; Raye, 1976). For instance, Lindsay and Johnson (1989) found that participants were more likely to make source confusions (i.e., to misattribute items that had been read as items that were seen) when they completed an old–new recognition test than when they completed a source-monitoring test. This discrepancy has implications for the process dissociation procedure, because the inclusion test is essentially an old–new recognition test and the exclusion test is a list-discrimination or source-monitoring test. Lindsay and Johnson suggested that criterial differences between the recognition and source-monitoring tests account for the different probabilities of making source confusions. Typical recognition tests, including the inclusion test, query participants about whether the target item was in a prior study episode and can be answered according to the item's degree of familiarity; the most familiar items indicate that the item was recently studied. List-discrimination tests and the exclusion test direct participants to distinguish between items from different study episodes, such as instructing participants to only call items old that were heard previously. Presumably, list-discrimination tests orient participants to recollect specific source-identifying information about an item, because familiarity cannot identify a particular study phase. In terms of the inclusion and exclusion tests, one would expect a greater reliance on recollective processes when participants are directed to make source judgments on the exclusion test than when they make old–new judgments on the inclusion test. Similarly, one would expect a greater reliance on familiarity for the inclusion task than for the exclusion task (e.g., see Lindsay & Johnson, 1989; Raye, 1976). Graf and Komatsu (1994) have expressed similar thoughts about the consistency assumption.

These reservations about the consistency assumption raise questions about two additional assumptions of this framework. First, it is assumed that familiarity's contribution to performance is automatic. Second, it is assumed that the process of recollection is all-or-none. Participants either correctly recollect the source of a studied item (e.g., that it was heard earlier) or they do not recollect anything about the studied item. For instance, Equation 2 states that the false recognition of an item studied in Phase 1 (e.g., anagram) on the exclusion test can only be a result of familiarity acting in the absence of recollection. In other words, the process of recollection can never produce errors, such as the misrecollection of an anagram item as a heard item. For the process-dissociation framework to work, it is necessary to assume that recollection is always accurate in order to establish the contrasting test situations that are essential for estimating the separate contributions of recollection and familiarity. When items are misrecollected, the processes of recollection and familiarity are no longer working in opposition on the exclusion test.

The following experiments examine these assumptions about familiarity and recollection. Specifically, in Experiment 1 we investigate the assumption that familiarity's influence is automatic, and in Experiment 2 we examine the consistency assumption and the assumption that recollection is an all-or-none process.

Experiment 1

To investigate the influence of familiarity, we constructed a situation in which the automatic influence of familiarity would compel a response that is opposite to the controlled use of familiarity. We used a task that has been used in a number of studies by Jacoby and colleagues (Jacoby, 1991; Jennings & Jacoby, 1993). In Phase 1 participants read words and solved anagrams, and in Phase 2 participants heard words. Two proportion conditions were constructed by varying the number of test items that were heard, read, and solved earlier as anagrams: In the *two-thirds-heard* condition, 60 test items were heard earlier, 15 test items were read earlier, and 15 were seen earlier as anagrams; in the *one-third-heard* condition, 30 items were heard earlier, 30 items were read, and 30 items were seen as anagrams. Notice that we do not change the overall proportion of old-to-new items in the two proportion conditions. In both proportion conditions, there were 90 studied items and 30 new items on the tests. On the inclusion test, participants were instructed to call all previously studied items old, and on the exclusion test they were instructed to say old about only the earlier heard items.

If participants control the use of familiarity, then its influence on the exclusion task should vary with changes in the proportion of heard items to other studied items (cf. Cheesman & Merikle, 1986; Neely, 1977). Participants may attempt to match their responses on this test to their perception of the proportion of heard items to the other studied items on the test. Participants should be less likely to call an item old on the exclusion task (i.e., saying it was heard) when one third of the studied items on the test were earlier heard than when two thirds of these items were earlier heard. In other words, there should be fewer false alarms for the anagrams and for the read items in the one-third-proportion condition than in the two-thirds-proportion condition. However, changes in the proportion of heard items to other studied items should not affect the role of familiarity on the inclusion test. This pattern of results on the exclusion and inclusion tests would result in differences in the estimated value of familiarity calculated for the two proportion conditions. On the other hand, if the influence of familiarity is automatic and unintentional then the particular proportion of heard items to other items should have no effect on either the false-alarm rate for the anagrams and read items or on the estimated value of familiarity.

Furthermore, we predicted that participants would be more likely to use familiarity strategically when attention is full at test than when attention is divided at test. If so, the two proportion conditions should yield similar false-alarm rates for the words read and seen earlier as anagrams on the exclusion test and similar estimated values of familiarity when attention is divided at test, but not when it is full. This result would strongly suggest that familiarity's influence is not automatic, but that it is controlled and attention demanding.

Method

There were three between-groups variables: the proportion of heard words to the other studied words on the test (i.e., either two thirds or one third), the type of test instruction (i.e., either inclusion or exclusion), and the manipulation of attention (i.e., either full or divided).

Participants. There were 120 paid student volunteers from Princeton University (7–9 per condition) and from the summer session of the University of California at Berkeley (6–8 per condition). Each was randomly assigned to one of the eight conditions.

Materials. The procedure and stimuli follow Jacoby (1991). The stimuli were 200 five-letter nouns. Three sets of 40 words were selected, each with an average word frequency of approximately 100 (Kučera & Francis, 1967). The sets of words were rotated so that each served as anagrams to be solved or words to be read in Phase 1 and as new words on the recognition test in Phase 3. The anagrams were presented with the second and fourth letters underlined and in their correct position. The remaining three letters in the anagram were randomly rearranged with the constraint that they did not all appear in their correct position (e.g., uasce). Each anagram had one unique solution. In Phase 1, 30 anagrams and 30 words to be read were randomly mixed with the constraint that no more than 3 items of either kind appeared consecutively. There were an additional 10 anagrams and 10 words to be read that were randomly intermixed and assigned to the first 10 and last 10 positions. These items were buffers and were not tested later. In Phase 2, a different set of 80 words was presented aurally to participants; the first 10 and last 10 items were buffers.

In Phase 3, the test lists contained different proportions of items that had been heard relative to the other studied items, although there was the same overall number of items (120) on the tests in both conditions. As in Jacoby (1991), the test list in the two-thirds-proportion condition included 60 items that had been heard, 15 items that had been presented as anagrams, 15 items that had been read, and 30 new items. In the one-third-proportion condition, the test list included 30 heard items, 30 anagram items, 30 read items, and 30 new items. We only scored the 15 anagram items, 15 read items, and the 15 new items that were common to both proportion conditions. However, we scored all 60 of the heard items in the two-thirds-proportion condition, and all 30 of the heard items in the one-third-proportion condition. In both conditions, 15 of the 30 new items were fillers and were presented in the initial 15 positions of the test list. These filler items were not scored and allowed the participants in the divided-attention conditions to adjust to the secondary listening task before receiving the critical items on the test list. All stimuli were presented in lowercase letters and in the center of the screen of a Macintosh SE computer (Apple, Inc., Cupertino, CA) with a letter size of approximately 5 mm × 7 mm.

The listening task in the divided-attention conditions was originally used by Craik (1982) and was constructed in the same manner as in Jacoby's (1991) studies. While taking the recognition test, participants simultaneously listened to a tape-recorded list of single-digit numbers; the numbers occurred at a 1.5-s rate. Participants were told to monitor the list for target sequences of 3 odd numbers in a row (e.g., 1, 7, 5). The list of numbers was randomly ordered with the restriction that 43 target sequences of 3 odd numbers occur within a set of 224 digits. There was the additional requirement that not more than 5 numbers and no fewer than 1 number could separate the target sequences.

Procedure. In Phase 1, participants were told that the experiment was about the processes involved in solving anagrams and reading words. Accordingly, they were told that whenever a word was presented in its normal form they should say it aloud and as quickly as possible. When the word was presented as an anagram, participants were informed that the second and fourth letters would be underlined and in their correct positions. Participants were told that they would have 30 s to rearrange the three letters that were not underlined so as to solve the anagram. The experimenter revealed the solution when this time passed and the participant had not solved the anagram. Participants were informed that the time to say each word aloud and to solve each anagram was being recorded. No mention was made of a later recognition test. Participants were told in Phase 2 that they would hear a series of words and that they should say aloud each word and remember them for a later memory test. The words were presented at a 2-s rate using a tape recorder.

Finally, participants were randomly assigned to one of eight different test conditions, produced by the three between-subjects variables of proportion of targets to nontargets (two thirds or one third), the manipulation of attention (full or divided), and the type of test (inclusion or exclusion). Participants receiving the inclusion test were instructed to call a word old if it had been encountered either as an anagram, as a word that was read, or as a word that was heard. This was essentially an old–new recognition test. For the exclusion test, participants were instructed to call a word old only if they remembered having heard the word. It was emphasized to participants that they would encounter words on the test that were from the first phase and from the second phase but that only the words from the second phase should be called old. In the full-attention conditions, participants completed only the recognition test, whereas in the divided-attention conditions participants received a secondary task while simultaneously completing the recognition test. The participants in the divided-attention condition were told that they would hear a series of single-digit numbers and that they should say now whenever they detected a target sequence of three odd numbers in a row. Participants were further informed that the experimenter would be monitoring their performance on the listening task and would say miss if they missed a target sequence. It was emphasized that it was important not to miss any target sequences and that the memory task should be completed automatically so as not to miss any target sequences on the listening task. The experimenter stopped the listening task once the recognition test had been completed.

In all conditions, the words on the test appeared a couple of double-spaced lines above the question "Old or New?" Participants were instructed to push the *l* key if the word was old and to push the *a* key if the word was new. After a recognition decision the screen cleared and was followed by a 1-s delay before presentation of the next test word.

Results

The significance level for all tests was set at $p < .05$. Table 1 shows the uncorrected and corrected probabilities of calling test items old in the inclusion and exclusion conditions.

False recognition of new items. There were different biases in the baseline rate of calling new items old on the two tests as a result of the different proportion and attention conditions. Collapsed across inclusion and exclusion tests, participants were more likely to falsely recognize new words in the two-thirds-proportion condition when attention was divided than when attention was full (.33 vs. .18), whereas in the one-third-proportion condition participants were somewhat less likely to recognize new words when attention was divided than when it was full (.23 vs. .29). A 2 × 2 × 2 analysis of variance (ANOVA; with factors of Proportion, Test, and Attention) of the false-alarm rate to the new items confirmed this pattern of bias with a significant interaction between Proportion and Attention, $F(1, 112) = 8.16$, $MSE = .04$. There were no other reliable effects in this analysis.

Inclusion test results. Following Jacoby (1991), the uncorrected data were corrected by subtracting the probability of falsely recognizing new words from the probability of correctly recognizing studied words (i.e., hits minus false alarms); the following analyses were conducted on the corrected recognition data for the anagram and read items, displayed inside the parentheses in Table 1. We separately analyzed the two Proportion conditions, in order to verify that we replicated Jacoby's (1991) results in the two-thirds-heard Proportion condition. In the two-thirds-heard Proportion condition, the corrected recognition rate of the read words was not affected by the absence or presence of Attention (.16 vs. .17), whereas participants were less likely to recognize words solved earlier as anagrams when attention was divided at Test than when it was full (.41 vs. .61). A 2 (Attention) × 2 (Type of Item) ANOVA of the corrected scores from the two-thirds-heard Proportion condition showed a reliable interaction between Attention and Type of Item, $F(1, 28) = 4.61$, $MSE = .03$. This result replicates Jacoby.

There was a different pattern of results in the one-third-heard proportion condition. The presence of our secondary task did not affect the corrected recognition rates of either the read words or the words presented as anagrams. A 2 (Attention) × 2 (Type of Item) ANOVA on the corrected scores from the one-third-Proportion condition showed no interaction between the Type of Item and the manipulation of Attention, although there was a main effect of item type, $F(1, 28) = 88.29$, $MSE = .03$. Participants recognized the words seen earlier as anagrams much more frequently than they recognized the words read earlier.

Exclusion test results. Table 1 also shows that, as predicted, manipulating the proportion of heard items on the exclusion test affected the corrected false-recognition rate of the anagram and read items when attention was full at test but not when attention was divided. The corrected false-recognition rate of the anagram and read items were analyzed with a 2 (Proportion) × 2 (Attention) × 2 (type of item) ANOVA. There was a significant interaction between Attention and Proportion, $F(1, 56) = 4.91$, $MSE = .07$, qualifying a main effect of Attention, $F(1, 56) = 6.42$, and a marginally significant effect of Proportion, $F(1, 56) = 3.06$, $p < .10$. There were no other reliable main effects or interactions. As shown in Table 1, when attention was full,

Table 1
Probability (P) of Calling a Word Old (and Corrected Probability [CP]) on the Inclusion and Exclusion Tests in Experiment 1

		Type of studied item						
		Anagram		Read		Heard		
Attention	Proportion heard	*P*	*CP*	*P*	*CP*	*P*	*CP*	New
		Inclusion test						
Full	Two thirds	.80	.61	.36	.17	.73	.54	.19
	One third	.85	.58	.42	.15	.72	.45	.27
Divided	Two thirds	.80	.41	.55	.16	.69	.30	.39
	One third	.79	.57	.39	.17	.65	.43	.22
		Exclusion test						
Full	Two thirds	.29	.11	.30	.12	.71	.53	.18
	One third	.21	−.10	.27	−.04	.64	.33	.31
Divided	Two thirds	.43	.17	.36	.10	.60	.34	.26
	One third	.38	.15	.40	.17	.55	.32	.23

Note. The two-thirds and one-third conditions refer to the proportion of studied items on the test that were heard earlier.

participants were less likely to falsely recognize Phase 1 words (i.e., words initially read and seen as anagrams) as heard words in the one-third-(−.07) than in the two-thirds-Proportion (.12) conditions, $F(1, 56) = 7.74$. However, there was no significant effect of Proportion on the false-recognition rate when attention was divided (.16 vs. .14), $F(1, 56) = .19$. The greater impact of Proportion under full rather than divided attention is exactly what was predicted and supports our belief that participants have some control over their use of familiarity on recognition tests. When attention is full, participants can strategically use familiarity by assessing the probability that a familiar item was heard and responding accordingly. Consequently, participants more often misattribute anagram and read items as heard items in the two-thirds-heard condition than in the one-third-heard condition. However, when attention is divided, participants are less able to use familiarity strategically and, thus, there is less of a difference between the two Proportion conditions in the rate at which the anagram and read items are misattributed as heard items.

Recognition of heard words. A 2 × 2 × 2 ANOVA (with factors of Proportion, Attention, and Test) on the recognition rate of the items that were heard earlier showed that participants were more likely to recognize items heard earlier when they completed the test with full attention (.70) than with divided attention (.62), $F(1, 112) = 6.73$, $MSE = .03$. This is not surprising, because many studies have shown that recognition performance is worse when the test is completed under conditions of divided rather than full attention (e.g., Jacoby, 1991; Park, Smith, Dudley, & Lafronza, 1989).

Of more interest is the fact that participants were less likely to recognize heard words when they completed the exclusion test (.63) than when they completed the inclusion test (.70), $F(1, 112) = 6.00$, $MSE = .03$. This result is surprising because participants were directed to call the heard words old on both the inclusion and exclusion tests. In other words, the task should be identical for the heard words on these two tests, and performance should be equivalent. However, nearly every other researcher using this procedure has found that the heard words are less likely to be recognized on the exclusion test than on the inclusion test (Hertel & Milan, 1993; Jacoby, 1991; Jacoby & Kelley, 1992; Jennings & Jacoby, 1993; Verfaellie & Treadwell, 1993). This pattern has been explained as a result of the influence of false recollection, that is, the influence of source confusions. Participants confused the words heard earlier with the words seen in Phase 1, and this hurt recognition performance on the exclusion test but not on the inclusion test. Because participants were instructed to exclude the items seen in Phase 1 on the exclusion test, the heard words that were misrecollected as Phase 1 words were incorrectly excluded. These source confusions did not affect recognition performance on the inclusion test because participants were instructed to call all of the previously studied items old. As discussed later, the presence of these source confusions suggests that recollection's influence is not equivalent on the inclusion and exclusion tests, in violation of the consistency assumption.

Results from the process-dissociation procedure. Recently, Jacoby and colleagues (Jacoby, Toth, & Yonelinas, 1993; Yonelinas, 1994; Yonelinas, Regehr, & Jacoby, 1995) proposed that differing response criteria between conditions or between the inclusion and exclusion tests, as measured by the false-alarm rate to new items, can contaminate the estimated values of familiarity. Jacoby and colleagues argued that, because the false alarms to the new items are only a result of the influence of familiarity, only familiarity should be corrected (rather than adjusting both familiarity and recollection). In addition, familiarity reflects a signal-detection process and should be measured in terms of d' rather than in terms of probabilities. Jacoby and colleagues noted that if the false-alarm rates to the new items do not differ between the inclusion and exclusion tests, then the d' scores for familiarity can be determined from standard d' tables by using the value of F as the hit rate and the proportion of false alarms as the false-alarm rate (see Yonelinas et al., 1995, for a thorough discussion of this issue). We followed this suggestion (however, see Roediger & McDermott, 1994, and Verfaellie, 1994, for further discussion of the issue of baseline response rates).

Because different participants completed the inclusion and exclusion tests, it was not possible to compute values of R and F d' from a single participant's responses. These values were computed for "macroparticipants" by combining the responses of 2 participants who completed identical test lists, one with the inclusion instructions and the other with the exclusion instructions. With 15 participants in each of eight between-groups conditions, we computed 15 values of R and F d' for each proportion condition under full and divided attention. The values of F d' were computed by treating the values of F as hit rates and treating the proportion of false alarms to the new items as false-alarm rates. Jacoby and colleagues (1993, p. 144) note that the estimate of automatic processes will be underestimated when participants score perfectly on the exclusion test (i.e., participants exclude all of the anagram and read items), producing an estimate of zero for the value of F (see Equation 2). They dealt with this problem by excluding Fs with values of zero. Similarly, we excluded the scores of 8 out of 60 participants from the following analyses.[2] The value of R is determined by subtracting the probability of calling an item old on the exclusion test from the probability of calling an item old on the inclusion test.

As seen in the top half of Table 2, the mean values of F d' for the anagram and read items paralleled the corrected false-recognition scores for these items on the exclusion test. The analysis of F d' scores is not simply redundant with the false-recognition scores because it shows that familiarity, defined by the process dissociation framework, can be used strategically. A 2 × 2 × 2 ANOVA (with factors of Proportion, Attention, and Item Type) on these scores indicated that the predicted interaction between proportion and attention was significant, $F(1, 48) = 4.25$, $MSE = .62$. The

[2] We obtained the same pattern of results when we did not exclude these participants. Moreover, the same pattern of results occur when we analyze the corrected F values (i.e., *M*).

Table 2
Estimated Values of Familiarity (F d′) and Recollection in Experiment 1

Attention	Proportion heard	Type of studied item: Anagram	Read	*M*
	Familiarity (F *d′*) process			
Full	Two thirds	1.50	0.58	1.04
	One third	0.94	0.13	0.54
Divided	Two thirds	1.21	0.40	0.81
	One third	1.34	0.54	0.94
	Recollection process			
Full	Two thirds	0.51	.06	.29
	One third	0.64	.15	.40
Divided	Two thirds	0.37	.19	.28
	One third	0.41	−.01	.20

Note: The two-thirds and one-third conditions refer to the proportion of studied items on the tests that were heard earlier.

interaction between Proportion and Attention is seen in the right-hand column of Table 2, containing the F d' scores that are averaged across the type of item. When the tests were taken with full attention, the value of F d' was reliably smaller in the one-third (.54) than in the two-thirds (1.04) condition, $F(1, 48) = 4.84$. However, there was no reliable effect of manipulating Proportion when Attention was divided at Test (.94 vs. .81), $F(1, 48) < 1$. If the influence of familiarity is automatic then there should be no effect of Proportion; the values of F d' should be the same in both Proportion conditions. Moreover, completing the recognition tests under conditions of full or divided attention should have no effect on the influence of familiarity if it is automatic. Thus, the data suggest that using familiarity is an attention-demanding, controlled process.

The ANOVA on the F d' scores also yielded a main effect of Item Type and no other significant main effects or interactions. The main effect of Item Type confirmed that the value of F d' was larger for the words that had been studied as anagrams than for those that had been read (1.25 vs. .42), $F(1, 48) = 34.80$, $MSE = .51$. Jacoby (1991) and Jennings and Jacoby (1993) also found that estimated familiarity was larger or equal for the words seen earlier as anagrams than for the words read earlier.

The lower half of Table 2 presents the mean values of R. The value of R was larger for the words seen earlier as anagrams (.48) than for the words read earlier (.10). A 2 × 2 × 2 ANOVA (with factors of Proportion, Attention, and Item Type) on the recollection scores revealed a main effect of Item Type, $F(1, 56) = 71.83$, $MSE = .06$. This result has been obtained in other studies by Jacoby and colleagues (e.g., Jacoby, 1991; Jennings & Jacoby, 1993) and suggests that there is more recollective information associated with the words solved as anagrams than with the read words. This analysis also yielded a marginally significant interaction between Item Type and Attention, $F(1, 56) = 3.82$, $p < .07$; there were no other reliable main effects or interactions. When collapsed across Proportion, in comparison with the value of R obtained with full attention, dividing attention decreased the value of R for the words seen as anagrams (.58 vs. .39) but did not affect the value of R for the read words (.11 vs. .09). This interaction suggests that the process of recollection that is associated with memory for the read items has different characteristics, such as not being vulnerable to disruptions of attention, than the process of recollection that is associated with memory for the anagram items. Notice that this difference in the values of R for the read and anagram items seems to contradict the notion that recollection is a controlled process whose influence should always decrease whenever control is impaired (i.e., with attention divided). However, with the present data, it is difficult to disentangle this interpretation from a possible floor effect for the read words, in that R would have been lower for the read items under divided attention conditions had the overall values of R been higher for these items under full attention.

Discussion

The results from the exclusion conditions confirmed our predictions: When attention was full at test, participants were less likely to confuse the Phase 1 items (anagram and read words) with the heard words when they received tests with the one-third-heard proportion than when they received tests with the two-thirds-heard proportion. When attention was divided at Test, there was no significant effect of Proportion on the false-alarm rates for the anagram and read items. Furthermore, the pattern of results from the values of F d' paralleled the results from the exclusion test. When attention was full, there was a greater effect of Proportion on the value of F d' than when attention was divided. The false-recognition rates and the values of F d' both suggest that, in the absence of recollecting the identity of an item, participants can strategically use familiarity to infer its identity when they are able to perceive the proportion of heard items to the other studied items on the test. Participants were less likely to infer that an item was heard in the one-third- than in the two-thirds-Proportion conditions and, consequently, made fewer errors of calling items heard that actually had been read or seen as an anagram. This strategy reduced the values of F d' in the one-third- as compared to the two-thirds-Proportion conditions because the values of F d' depend on the false-alarm rates from the exclusion test. In addition, this ability to strategically use familiarity is contingent on the presence of attention. In short, these results contradict the assumption of the process-dissociation framework that familiarity's influence is automatic. Consequently, there is no way to guarantee that the role of familiarity is equivalent on the inclusion and exclusion tests, which undermines the logic of the process-dissociation method.

In addition, our results suggest that the role of recollection is not equivalent in the two tests, violating the consistency assumption that both recollection and familiarity remain unchanged in the inclusion and exclusion tests. The heard items were less likely to be recognized on the exclusion test than on the inclusion test. This pattern of results has occurred in nearly every other study that has used the process-dissociation procedure with recognition tests (Hertel & Milan, 1993; Jacoby, 1991; Jacoby & Kelley, 1992; Jennings & Jacoby, 1993; Verfaellie & Treadwell, 1993). These researchers explained this pattern as the influence of false recollection, that is, the influence of source confusions. Participants confused the words heard earlier with the words that were read or seen earlier as an anagram. With the exclusion instructions, the heard words that were falsely recollected as anagram or read items were excluded as well. These source confusions did not affect the rate at which the heard items were recognized on the inclusion test, because participants were instructed to call all of the previously studied items old. Therefore, the misrecollection of heard words can lower the recognition rate of the heard items on the exclusion test but not on the inclusion test. Experiment 2 demonstrates that the occurrence of misrecollection causes an imbalance in the role of recollection on the inclusion and exclusion tests, in violation of the consistency assumption.

Experiment 2

To demonstrate that misrecollecting heard items causes an imbalance in the role of recollection on the inclusion and exclusion tests, consider the following equations: According to Equation 1, a correct recognition of a previously studied item, such as a heard item, on the inclusion test can be a result of either recollecting the item or a result of familiarity, in the absence of recollection.

Although Jacoby and colleagues (e.g., Jacoby, 1991; Jennings & Jacoby, 1993) have not formally expressed Equation 3, it follows from the logic of the process dissociation procedure that a correct recognition of a heard item on the exclusion test (remember that on this test, participants are instructed to only call the heard items old) must also be a result of either recollection or familiarity in the absence of recollection:

$$p(\text{correct recognition} \mid \text{exclusion test}) = \mathrm{R} + \mathrm{F}(1 - \mathrm{R}). \quad (3)$$

Thus, it follows that the probability of correctly recognizing heard items on the inclusion test should equal the probability of correctly recognizing heard items on the exclusion test. Whenever the recognition rates of the heard items are different on these two tests then the consistency assumption has been violated because the influence of either familiarity or recollection (or both) is not the same on the two tests. We agree with Jacoby and colleagues (e.g., Jennings & Jacoby, 1993) that misrecollection of the heard items probably accounts for the lower recognition rate of these items on the exclusion test than on the inclusion test. This means that the process of recollecting heard items is less influential on the exclusion test than on the inclusion test.

However, if the heard items can by misrecollected as other studied items, it is plausible that the anagram or read items can be misrecollected as heard items. If misrecollection can cause participants incorrectly to call an anagram item old on the exclusion test, then this contradicts the assumption that the false recognition of an item on the exclusion test can be a result only of the influence of familiarity in the absence of recollection. This assumption is represented by Equation 2.

The process-dissociation framework assumes that whenever participants recollect a studied item, they either correctly recollect the source of the studied item (e.g., that it was heard earlier) or they recollect nothing about the source of the item; recollection is all-or-none.

According to the source-monitoring framework of Johnson and colleagues (e.g., Johnson, et al., 1993), the phenomenal experience of recollection does not always yield accurate source identification. One factor that affects how likely a memory will be misrecollected as another memory is the similarity between the two memories. This framework assumes that the attribution of a memory to a source is a decision-making activity that typically involves evaluating the characteristics of the memory in light of currently activated agendas (Johnson & Hirst, 1993). A number of studies have shown that memories containing similar characteristics are more likely to be confused with each other than memories containing different characteristics (e.g., Ferguson, Hashtroudi, & Johnson, 1992; Foley, Johnson, & Raye, 1983; Hashtroudi, Johnson, & Chrosniak, 1989; Johnson, Foley, & Leach, 1988; Lindsay & Johnson, 1991; Lindsay, Johnson, & Kwon, 1991).

Experiment 2 manipulated the similarity of the target items to the other studied nontarget items to show when the target items were more or less likely to be misrecollected as a nontarget. By *target,* we refer to the items that should be called old on the exclusion test. For example, in Experiment 1 the targets were the heard items. The prediction is that the recognition rate of the targets will be lower on the exclusion test than on the inclusion test when the target and nontarget items are similar as compared to when they are dissimilar. Misrecollections are important because they illustrate that recollections do not always yield a correct response. Experiment 2 uses the same overall procedure that was used in Experiment 1. As shown in the Appendix, in Phase 1 participants read words and solved anagrams. In Phase 2, half of the participants heard words, and the remaining participants solved word fragments. The word fragments were created by removing two letters (e.g., *sw__d).* The processes involved in solving anagrams are very similar to the processes used to solve word fragments. The anagrams are less similar to the heard words than they are to the words presented as word fragments. Consequently, it is more likely that the items seen as anagrams will be misrecollected as word fragments than that they will be misrecollected as heard items.

As shown in the Appendix, in Phase 3 of Experiment 2 there were four different tests: the standard inclusion test as well as three different forms of the exclusion test. Some of the participants who heard words in Phase 2 received what is referred to as the exclusion(heard+) test; these participants were directed to call the heard words old. Some of the participants who solved word fragments in Phase 2 received what is referred to as the exclusion(fragment+) test; they were instructed to call the words seen as word fragments old. Finally, some participants received an exclusion(anagram+) test and were instructed to call only the words presented as anagrams old. Note that the exclusion(heard+) test in Experiment 2 corresponds to the standard exclusion test in Experiment 1 and in studies by Jacoby and colleagues (e.g., Jacoby, 1991; Jacoby & Kelley, 1992).

Just as Equations 1 and 3 indicate that both recollection and familiarity contribute to the correct recognition of heard items on the inclusion and exclusion tests (this also applies to the recognition rate of the items seen as word fragments), these same equations must account for the correct recognition rate of the items seen as anagrams on the inclusion and exclusion(anagram+) tests. If there are different recognition rates for the anagram items on the inclusion and exclusion(anagram+) tests, recollection, familiarity, or both are having different influences on the inclusion and exclusion tests.

When participants have solved word fragments in Phase 2, we expect the recognition rate for the items seen as anagrams to be lower on the exclusion(anagram+) test than on the inclusion test. Whenever a word seen as an anagram is misrecollected as a word fragment, this will lower the recognition rate of the anagrams on the exclusion(anagram+) test but not on the inclusion test. Thus, when participants have solved word fragments, there should be a larger difference in the recognition rate of the anagrams on the exclusion(anagram+) and inclusion tests than when participants have heard words. In addition, there should be a larger false-recognition rate of the anagram items on the exclusion(fragment+) test than on the exclusion(heard+) test because of the greater likelihood that the words seen as anagrams will be misrecollected as words solved as word fragments than as heard words.

Method

Participants. There were 70 Princeton students in Experiment 2. Each was paid to participate and was randomly assigned to one of six conditions (see the *Procedure* section) for a total of 11 to 12 participants in each condition.

Materials. A set of 135 words similar to those used in Experiment 1 served as the stimuli. Three sets of 20 words were rotated so that each served as anagrams to be solved, words to be read in Phase 1, and new words on the recognition test in Phase 3. The anagrams were constructed in the same manner as in Experiment 1. Participants were presented with a randomly mixed list of 20 anagrams and 20 words to be read; no more than 3 items of each kind appeared consecutively.

In Phase 2 a different set of 60 words was presented aurally to half of the participants and as word fragments to the remaining participants. The word fragments were constructed by showing three letters and underlining the spaces of the two missing letters (e.g., *sw__d)*. The missing letters were chosen so that there was only one solution possible.

The test lists in Phase 3 were constructed in an identical manner as the ones used in the two-thirds-Proportion condition of Experiment 1. Each test included 120 words: 60 words that were either heard or presented as a word fragment (depending on the condition), 15 words that had been read, 15 words that had been presented as anagrams, and 30 words that were new.

Procedure. The procedure for Phase 1 was identical to Phase 1 in Experiment 1. In Phase 2, half of the participants were told that they would hear a series of words and that they should say aloud each word and remember it for a later memory test. The words were presented at a 2-s rate using a tape recorder. The remaining participants were told that they would see a series of word fragments, each having one noun as a solution. Participants were informed that they would have 30 s to solve and say aloud the solution to the fragment. The experimenter revealed the solution to the fragment when this time passed. They were told that their memory for the solutions to the fragments would be tested later.

Finally, the six conditions were as follows. The participants who heard words in Phase 2 completed either an inclusion test, an exclusion(anagram+) test, or an exclusion(heard+) test. On the inclusion test, participants were told to call all of the previously studied words old (i.e., the anagram, read, and heard words). Participants receiving the exclusion(anagram+) test were instructed to call a word old only if they remembered having seen the word presented as an anagram. On the exclusion(heard+) test, which was identical to the exclusion test in Experiment 1, participants were instructed to call words old only if they remembered having heard this word. The participants who solved word fragments in Phase 2 received either an inclusion test, an exclusion(anagram+) test, or an exclusion(fragment+) test. The instructions for the inclusion and exclusion(anagram+) tests are identical to those given to participants who heard words in Phase 2. However, for the exclusion(fragment+) test, participants were instructed to call words old only if they remembered having seen the item presented as a word fragment.

Results and Discussion

The significance level for all tests was set at $p < .05$.

Recognition rate of the words solved earlier as anagrams on the inclusion and exclusion(anagram+) tests. In both the inclusion test and the exclusion(anagram+) test, participants must call the words seen as anagrams old. If recollection and familiarity are having the same influence on both tests then there should be no difference in the recognition rate of the anagrams on the two tests. A 2 × 2 ANOVA, with factors of Phase 2 Activity (i.e., heard-fragment) and Type of Test, that is, inclusion- exclusion(anagram+), was performed on the recognition rate of the items seen as anagrams. As shown in Table 3, there was a significant main effect of Phase 2 Activity, $F(1, 42)$ =

Table 3
Recognition Rate of the Anagrams on the Inclusion and Exclusion(Anagram+) Tests, the False-Alarm Rate for the New Words, and the Corrected Recognition Rate of the Anagrams in Experiment 2

Test	Recognition	False alarm	Corrected recognition
Heard words in Phase 2			
Inclusion	.81	.23	.58
Exclusion(anagram+)	.89	.08	.81
Solved word fragments in Phase 2			
Inclusion	.74	.10	.64
Exclusion(anagram+)	.52	.12	.40

25.75, MSE = .02. Collapsed across Test, the anagram items were more likely to be recognized when participants heard words (.85) in Phase 2 than when they solved word fragments (.63). As predicted, there was a significant interaction between Phase 2 Activity and the Type of Test, $F(1, 42) = 12.00$, MSE = .02. When participants solved word fragments in Phase 2, they were less likely to recognize words seen as anagrams on the exclusion(anagram+) test (.52) than on the inclusion test (.74), $F(1, 42) = 12.10$. There was no reliable difference in the recognition rate of the words seen as anagrams when participants heard words in Phase 2, .89 on the exclusion(anagram+) test versus .81 on the inclusion test.

As seen in Table 3, there was some variability in the false-alarm rate to the new items. A 2 (Phase 2 Activity) × 2 (Test Type) ANOVA on the false-alarm rate to the new items revealed a marginally significant interaction and no reliable main effects, $F(1, 42) = 3.53$, $MSE = .03$, $p < .07$. There was no difference in the false-alarm rates on the two tests when participants solved word fragments in Phase 2. However, when participants heard words in Phase 2, they were more likely to falsely recognize new items on the inclusion test (.23) than on the exclusion(anagram+) test (.08). Because of the possibility that participants had different response criteria on the two tests, we analyzed the corrected recognition rates of the anagram items on the inclusion and exclusion(anagram+) tests. A 2 (Phase 2 Activity) × 2 (Test Type) ANOVA revealed a reliable interaction, $F(1, 42) = 16.31$, $MSE = .04$, that qualifies the main effect of Phase 2 Activity. As predicted, when participants solved word fragments they were less likely to recognize the anagram items on the exclusion(anagram+) test than on the inclusion test (.40 vs. .64), $F(1, 42) = 7.80$. However, participants were more likely to recognize the anagram items on the exclusion(anagram+) test than on the inclusion test when participants heard words in Phase 2 (.81 vs. .58), $F(1, 42) = 8.37$. Interestingly, in a different paradigm, we have previously found that recognition rates for a class of items can be higher on a source-monitoring test, like the exclusion test, that orients participants to think about the source of the test item than on an old–new recognition test, like the inclusion test (e.g., Dodson & Johnson, 1993).

The recognition rates of the anagrams were different on the two tests, demonstrating that recollection, familiarity, or both were having inconsistent influences on the inclusion and exclusion tests. When participants solved word fragments in Phase 2, the anagram items were sometimes misrecollected as word fragments that caused participants to incorrectly call the anagram items new on the exclusion(anagram+) test, reducing the recognition rate for these items. The occurrence of this misrecollection did not affect the recognition rate of the anagram items on the inclusion test, because participants were instructed to call all of the previously studied items old. In short, just as Jacoby and others have found that the heard items are mistaken for other items on the exclusion(heard+) test, so we have found that the items seen as anagrams can be mistaken for word fragments on the exclusion(anagram+) test. However, the occurrence of such source confusions means that recollection is having different roles on the inclusion and exclusion tests, in violation of the consistency assumption, and therefore, invalidates using such data to derive values of recollection and familiarity.

False recognition of the anagram items on the exclusion(fragment+)–exclusion(heard+) tests. As seen in Table 4, participants were more likely to confuse anagrams with word fragments on the exclusion(fragment+) test (.45) than to confuse anagrams with heard items on the exclusion(heard+) test (.26), $F(1, 23) = 5.98$, MSE = .03. Because of the different baseline rates of calling new items old on these two exclusion tests, we also examined the corrected false-recognition rates. Using these corrected scores, participants were more than five times as likely to incorrectly call an item seen as an anagram old on the exclusion(fragment+) test than on the exclusion(heard+) test (.39 vs. .07, respectively), $F(1, 23) = 17.42$, MSE = .03. This suggests that memories of the words processed as anagrams are much more likely to be misrecollected as words that were processed as word fragments than as heard words.

General Discussion

The experiments reported here examined three assumptions of the process-dissociation procedure: First, that the influence of familiarity is automatic; second, that the influences of familiarity and recollection remain the same on the inclusion and exclusion tests; and third, that the process of recollection is all-or-none. Experiment 1 demonstrated that familiarity's influence is not automatic. Manipulating the proportion of heard words to the other studied words on the inclusion and exclusion tests affected the false-alarm rate of the words read and seen earlier as anagrams on the exclusion test as well as the derived values of F d'. These data suggest that participants can strategically use familiarity by assessing the probability that an item was heard and respond accordingly. When there was a high proportion of heard items on the exclusion test, participants were likely to infer that a familiar, but not recollected, test item was probably heard earlier (i.e., because participants were instructed to only call heard items old on this test). The high false-recognition rate for the anagram and read items increased the value of F d' for these items. Participants were less likely to infer that an item was heard when there was a low

Table 4
False Recognition Rate of the Anagrams on the Exclusion(Fragment+) and Exclusion(Heard+) Tests, the False-Alarm Rate for the New Words, and the Corrected False-Recognition Rate of the Anagrams in Experiment 2

Test	False recognition of anagrams	False alarms to new items	Corrected false recognition
Exclusion(fragment+)	.45	.06	.39
Exclusion(heard+)	.26	.19	.07

proportion of heard items on the test. Consequently, there was a low false-alarm rate and a low value of F d'. However, when the exclusion test was completed with divided attention, there was no significant effect of the proportion of heard items on the false-alarm rate of the words read or seen earlier as anagrams or on the values of F d' for these items. In short, the sensitivity of F d' and of the false-alarm data to the effects of proportionality and divided attention are taken as evidence that familiarity's influence is not automatic but is controlled and attention demanding.

Experiment 2 demonstrates that whenever there are different correct recognition rates of a class of items on the inclusion and exclusion tests, then the consistency assumption has been violated. Numerous researchers have found that the correct recognition rate of the heard items can be smaller on the exclusion test than on the inclusion test (our Experiment 1; Hertel & Milan, 1993; Jacoby, 1991; Jacoby & Kelley, 1992; Jennings & Jacoby, 1993; Verfaellie & Treadwell, 1993). We agree with Jacoby and colleagues that misrecollection or source confusion causes the different correct recognition rates of these items on the two tests. Experiment 2 confirmed this explanation by manipulating the similarity of the target and nontarget old items, a variable generally accepted as affecting source recollection, to show that misrecollection (or source confusion) caused the different correct recognition rates of target items on the inclusion and exclusion tests. Specifically, the recognition rate of words seen as anagrams was lower on the exclusion(anagram+) test (where participants must call only the anagram items old) than on the inclusion test when participants had solved word fragments in Phase 2. However, there was no difference in the recognition rate of the anagram items on these two tests when participants heard words in Phase 2. The misrecollection of the anagram items as word fragments decreased the recognition rate of the anagram items on the exclusion(anagram+) test but did not affect the recognition rate of the anagram items on the inclusion test. In other words, recollection does not have the same influence on the inclusion and exclusion tests. Although misidentifications of this kind are quite understandable from a source-monitoring perspective, they are a problem for the process-dissociation framework and invalidate the estimated values of recollection and familiarity. An essential part of the process-dissociation framework is the goal of constructing contrasting test situations so that familiarity and recollection work together to produce a response in the inclusion test and they work in opposition in the exclusion test. When participants misidentify items seen as anagrams as word fragments, familiarity and recollection are not working in opposition on the exclusion test, because both processes are contributing to the false recognition of items.

In addition, the occurrence of misrecollections violates the assumption that the process of recollection is all-or-none. According to the process-dissociation framework (represented by Equation 2), the false recognition of target items on the exclusion test can only occur when familiarity is acting in the absence of recollection; errors can never be a result of recollection. In other words, participants either recollect the source of an item or they do not recollect anything about the item. However, this assumption conflicts with the results of Experiment 2, which shows that participants can falsely recognize items because of misrecollection.

A different conception of recollection is provided by the source-monitoring framework (Johnson et al., 1993). According to this framework, the phenomenal experience of recollection does not always lead to a correct source identification. Instead of recollecting a tag that specifies the source of a memory, this framework assumes that the attribution of a memory to a source is a judgment that typically involves assessing the characteristics of a memory in light of a currently activated task or agenda (Johnson & Hirst, 1993). Furthermore, the characteristics of a memory will vary depending on the kind and amount of specific information that is recollected. For example, some memories may only contain a vague sense of familiarity and no specific information at all. Without the use of inferential strategies, such as those used in Experiment 1, these memories would only allow participants to discriminate these items from new words. Other memories for items may have some specificity, such as recollecting that the item was seen, and can be discriminated from other familiar items as long as the two are not too similar, such as distinguishing between the anagram and heard words. Still other memories are full of details, such as the appearance of particular letters, and support very specific discriminations, such as discriminating anagrams from word fragments. Describing memory in terms of various degrees of specific information counters the notion that there are two distinct forms of memory, either familiarity only or complete recollection of an event.

Our results have implications for the current debate over the relationship between the measures of familiarity and recollection (e.g., Jacoby, Toth, Yonelinas, & Debner, 1994; Joordens & Merikle, 1993). An article by Joordens and Merikle (1993) questioned the assumption of the process-dissociation framework that the conscious process of recollection and the automatic process of familiarity are independent and argued that whenever there is recollection there is also familiarity (i.e., familiarity and recollection are redundant). However, because we have shown that the purported measures of controlled and automatic processes (i.e., R and F) are not measuring what they seek to measure, it is questionable what either an independent or a redundant relationship means.

Buchner, Erdfelder, and Vaterrodt-Plünnecke (1995) recently extended the process-dissociation model by adding guessing parameters to remove response biases from the estimates of unconscious (familiarity) and conscious (recollection) processes. An analysis of the data from our Experiment 1 using the extended Buchner et al. multinomial model yielded comparable results to those we reported. Specifically, the estimate of familiarity (parameter u_c-) was significantly higher in the two-thirds-Proportion condition (.29) than in the one-third-Proportion condition (.15) when attention was full, $G^2(1) = 6.48$, $p < .05$, but not when attention was divided (.32 and .34 for two-thirds and one-thirds conditions, respectively), $G^2(1) = .05$.[3] Thus, even with Buchner et al.'s improved measure the influence of familiarity is not automatic in the sense required by the logic of the process-dissociation framework. It should also be noted that the extended model, like the process-dissociation framework, treats recollection as an all-or-none process. Although the extended model is preferable to the original process-dissociation model as a measurement tool when the assumptions of the process-dissociation approach hold, the extended model was not intended to address the limitations of the process-dissociation framework as a general theoretical account of recognition and source memory. Next, consider the implications of our results for other paradigms.

The process-dissociation procedure has been applied to a word-stem task that involves studying words (e.g., water) and then completing word stems (e.g., *wat__*) in a subsequent test phase (e.g., Jacoby et al., 1993). Inclusion instructions direct participants to complete the word stem with a previously studied item (i.e., *water),* or, failing that, with the first word that completes the stem. It is assumed that both automatic and controlled processes contribute to completing stems with studied words. Exclusion instructions direct participants to complete the stem with a word that had not been studied before (e.g., *watch).* If participants incorrectly complete a word stem with a studied item (e.g., *water),* then it is assumed that this is because the studied word automatically came to mind and was accepted as a completion because there was no recollected information (i.e., the controlled process) to oppose this judgment.

As with other paradigms, the consistency assumption must be correct in order for the process-dissociation procedure to derive valid estimates of the contributions of the automatic and controlled processes. On the basis of the automatic process alone, participants must accept studied words as completions to the same extent with the inclusion and exclusion instructions. However, Experiment 1 shows that in a recognition task participants do not automatically accept items in the absence of recollection, but rather they assess the word's familiarity and evaluate the probability that an item originated from a particular source. If participants assess the familiarity of a word in this word-stem paradigm (and in the absence of recollection), it seems unlikely that participants will accept studied words to the same extent with the inclusion and exclusion instructions, especially because participants are looking to accept studied words in the inclusion condition and to reject these words in the exclusion condition. It seems likely that participants will use different familiarity criteria with the inclusion and exclusion instructions. That is, participants may use a liberal familiarity criterion when they are instructed to remember studied items (inclusion instructions) because they are looking to accept any word that feels familiar. In contrast, participants may use a conservative familiarity criterion when they are instructed to reject studied items, that is, to exclude all items that feel familiar (exclusion instructions). The use of different familiarity criteria with the inclusion and exclusion instructions violates the consistency assumption and invalidates the derived values using the process-dissociation equations.

Our findings are also relevant to interpreting the results from other paradigms, such as the false-fame effect and the judgment of truth, that have been used to support the view that familiarity's influence is automatic (e.g., Begg, Anas, & Farinacci, 1992; Jacoby, Kelley, et al., 1989; Jacoby, Woloshyn, & Kelley, 1989). For example, Jacoby, Kelley, et al. (1989, Study 1) showed participants a list of nonfamous names, such as Sebastian Weisdorf. Subsequently, some participants received an immediate test and others received a test after a delay of 24 hr. Both groups of participants were told that the test included some of the previously read nonfamous names, new nonfamous names, and famous names and that their task was to judge whether or not the name was famous. In the immediate test condition, old nonfamous names were less likely to be judged famous than the new nonfamous names; in the delayed test condition, the old nonfamous names were more likely or equally likely to be judged famous than were the new nonfamous names.

When the test was immediate it was likely that participants recollected having read the old nonfamous names and thus did not mistake these names for famous names. Delaying the test apparently reduced the influence of conscious recollection but did not affect the familiarity of the earlier read nonfamous names. Jacoby, Kelley, et al. (1989, p. 327) concluded that the increase in mistakenly calling nonfamous names famous "must result from an unconscious influence of the past, because conscious recollection of the name as previously read in the list of nonfamous names would dictate an opposite response."

An alternative explanation of the false-fame effect in the previous experiment involves not the unconscious influence of familiarity but rather the conscious use of familiarity. When participants cannot recollect the source of an old nonfamous name on the test that feels familiar, they may realize that the name's familiarity could easily be a result of either the name's fame or its having been read earlier. If participants perceive that half of the names on the test in Jacoby, Kelley, et al.'s study (1989) were famous and the other half were nonfamous, then for half of the names that are familiar, but are not recollected, participants will guess that the name was famous. Because participants will recollect fewer names on the delayed test than on the immediate test, there will be a larger percentage of names on the delayed test than on the immediate test that participants will have to guess famous or nonfamous. This will increase the overall number of names that will be called famous on the delayed test as compared to the immediate test, resulting in a false-fame effect on the delayed test and not on the

[3] We would like to thank Ute Bayen for conducting these analyses and for helpful discussions of multinomial modeling approaches to source monitoring. The multinomial analyses were completed with a computer program by Hu (1993; see Hu & Batchelder, 1994). G^2 is a log-likelihood ratio statistic that is asymptotically chi-square distributed (see Batchelder & Riefer, 1990).

immediate test. This explanation does not rely on assuming that familiarity's influence is unconscious.

Conclusion

In sum, these experiments show that for understanding recognition memory, the process-dissociation framework is not achieving what it sought to accomplish, namely to measure an automatic influence of familiarity and a controlled influence of recollection. There are three important conclusions: First, familiarity's influence is not automatic; second, the influences of recollection and familiarity, in contrast to the consistency assumption, do not remain the same on the inclusion and exclusion tests; and third, recollection is not a perfect process but can result in varying types and amounts of specific detail and can produce source confusions when participants misrecollect an item of one type as an item of another type. By treating source confusions as an automatic consequence of the influence of familiarity, the process-dissociation framework ignores the active, decision-making aspect of remembering and the possibility that recollected information can produce source misidentifications as well as correct responses. As Experiment 1 suggests, participants may not recollect the source of their memory for an item but may still strategically infer the likely source, given their knowledge of the test environment. Participants can flexibly weight particular memory characteristics as a basis for a judgment, requiring more information in some circumstances and less information in others (e.g., Dodson & Johnson, 1993; Johnson et al., 1993; Lindsay & Johnson, 1989; Raye, 1976). As Experiment 2 suggests, whether the recollected information is sufficient for accurate performance on an exclusion test (i.e., a source-monitoring test) depends on the similarity between the target and the alternative items (Ferguson et al., 1992; Johnson et al., 1988).

References

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review, 79,* 97–123.

Anderson, J. R., & Bower, G. H. (1974). A propositional theory of recognition memory. *Memory & Cognition, 2,* 406–412.

Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (Vol. 1, pp. 243–293). San Francisco: W. H. Freeman.

Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review, 97,* 548–564.

Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General, 121,* 446–458.

Buchner, A., Erdfelder, E., & Vaterrodt-Plünnecke, B. (1995). Toward unbiased measurement of conscious and unconscious memory processes within the process dissociation framework. *Journal of Experimental Psychology: General, 123,* 137–160.

Cheesman, J., & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology, 40,* 343–367.

Craik, F. I. M. (1982). Selective changes in encoding as a function of reduced processing capacity. In F. Klix, J. Hoffman, & E. van der Meer (Eds.), *Cognitive research in psychology* (pp. 152–161). Berlin, Germany: Deutscher Verlag der Wissenschaften.

Craik, F. I. M., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 474–479.

Curran, T., & Hintzman, D. L. (1995). Violations of the independence assumption in process dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 531–547.

Dodson, C. S., & Johnson, M. K. (1993). The rate of false source attributions depends on how questions are asked. *American Journal of Psychology, 106,* 541–557.

Ferguson, S. A., Hashtroudi, S., & Johnson, M. K. (1992). Age differences in using source-relevant cues. *Psychology and Aging, 3,* 443–452.

Foley, M. A., Johnson, M. K., & Raye, C. L. (1983). Age-related changes in confusion between memories for thoughts and memories for speech. *Child Development, 54,* 51–60.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91,* 1–67.

Graf, P., & Komatsu, S. (1994). Process dissociation procedure: Handle with caution! *European Journal of Cognitive Psychology, 6,* 113–129.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108,* 356–388.

Hashtroudi, S., Johnson, M. K., & Chrosniak, L. D. (1989). Aging and source monitoring. *Psychology and Aging, 4,* 106–112.

Hertel, P. T., & Milan, S. (1993). *Depressive deficits in recognition: Dissociation of recollection and familiarity.* Poster session presented at the annual meeting of the Psychonomics Society, Washington, DC.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95,* 528–551.

Hintzman, D. L., & Curran, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory and Language, 33,* 1–18.

Hu, X. (1993). Statistical inference program for multinomial binary tree models (Version 1.0) [Computer program]. Irvine: University of California.

Hu, X., & Batchelder, W. H. (1994). The statistical analysis of general processing tree models with the EM algorithm. *Psychometrika, 59,* 21–47.

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.

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., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 3,* 306–340.

Jacoby, L. L., & Kelley, C. M. (1992). Unconscious influences of memory: Dissociations and automaticity. In D. Milner & M. Ruggs (Eds.), *Consciousness and cognition: Neuropsychological perspectives* (pp. 201–233). London: Academic Press.

Jacoby, L. L., Kelley, C. M., Brown, J., & Jasechko, J. (1989).

Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology, 56,* 326–338.
Jacoby, L. L., Lindsay, D. S., & Toth, J. P. (1992). Unconscious influences revealed: Attention, awareness, and control. *American Psychologist, 47,* 802–809.
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.
Jacoby, L. L., Toth, J. P., Yonelinas, A. P., & Debner, J. A. (1994). The relationship between conscious and unconscious influences: Independence or redundancy? *Journal of Experimental Psychology: General, 123,* 216–219.
Jacoby, L. L., Woloshyn, V., & Kelley, C. M. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General, 118,* 115–125.
Jennings, J. M., & Jacoby, L. L. (1993). Automatic versus intentional uses of memory: Aging, attention, and control. *Psychology and Aging, 8,* 283–293.
Johnson, M. K., Foley, M. A., & Leach, K. (1988). The consequences for memory of imagining in another person's voice. *Memory & Cognition, 16,* 337–342.
Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletine, 114,* 3–28.
Johnson, M. K., & Hirst, W. (1993). MEM: Memory subsystems as processes. In A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp. 241–286). Hove, England: Erlbaum.
Joordens, S., & Merikle, P. M. (1993). Independence or redundancy? Two models of conscious and unconscious influences. *Journal of Experimental Psychology: General, 122,* 462–467.
Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.
Lindsay, D. S., & Johnson, M. K. (1989). Eyewitness suggestibility and memory for source. *Memory & Cognition, 17,* 349–358.
Lindsay, D. S., & Johnson, M. K. (1991). Recognition memory and source monitoring. *Bulletin of the Psychonomic Society, 29,* 203–205.
Lindsay, D. S., Johnson, M. K., & Kwon, P. (1991). Developmental changes in memory source monitoring. *Journal of Experimental Child Psychology, 52,* 297–318.
Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87,* 252–271.
Murdock, B. B., Jr. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review, 89,* 609–626.
Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106,* 226–254.
Park, D. C., Smith, A. D., Dudley, W. N., & Lafronza, V. N. (1989). Effects of age and a divided attention task presented during encoding and retrieval on memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1185–1191.
Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing in cognition: The Loyola Symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.
Raye, C. L. (1976). Recognition: Frequency or organization? *American Journal of Psychology, 89,* 645–658.
Roediger, H. L., & McDermott, K. B. (1994). The problem of differing false-alarm rates for the process dissociation procedure: Comment on Verfaellie and Treadwell (1993). *Neuropsychology, 8,* 284–288.
Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing. I. Detection, search, and attention. *Psychological Review, 84,* 1–66.
Underwood, B. J. (1972). Word recognition memory and frequency information. *Journal of Experimental Psychology, 94,* 276–283.
Verfaellie, M. (1994). A re-examination of recognition memory in amnesia: Reply to Roediger and McDermott. *Neuropsychology, 8,* 289–292.
Verfaellie, M., & Treadwell, J. (1993). Status of recognition memory in amnesia. *Neuropsychology, 2,* 5–13.
Yonelinas, A. P. (1994). Receiver operating characteristics in recognition memory: Evidence for a dual process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1341–1354.
Yonelinas, A. P., Regehr, G., & Jacoby, L. L. (1995). Incorporating response bias in a dual-process theory of memory. *Journal of Memory and Language, 34,* 821–835.

Appendix

Design of Experiment 2

Phase 1: Participants read words and solve anagrams.
Phase 2: Participants either hear words or solve word fragments.
Phase 3: Participants complete one of four different tests:
- Inclusion—call all studied items *old.*
- Exclusion(heard+)—only call the heard items old.
- Exclusion(fragment+)—only call the items seen as word fragments old.
- Exclusion(anagram+)—only call items seen as anagrams old.

Received July 6, 1994
Revision received February 2, 1995
Accepted October 3, 1995 ■