Automatic Versus Intentional Uses of Memory: Aging, Attention, and Control

Janine M. Jennings and Larry L. Jacoby

In 2 experiments, the authors used a process dissociation procedure (Jacoby, 1991) to separately examine the effects of aging on automatic and consciously controlled memory processes. In Experiment 1, a group of young adults in either a full-attention or divided-attention condition were compared with a group of elderly adults on a fame judgment task. Both age and divided attention had a detrimental effect on consciously controlled memory processing but left automatic processing intact. In Experiment 2, the same age-related pattern was found using a more demanding forced-choice recognition paradigm.

There has been a great deal of recent interest in dissociations between performance on direct and indirect tests of memory found with the elderly (for reviews, see Craik & Jennings, 1992; Light, 1991). Direct tests of memory, which ask people to report on a past event, reveal pronounced declines in memory performance with age. The elderly are impaired, relative to the young, on tests of free recall, cued recall, and recognition (e.g., Craik, 1986; Craik & McDowd, 1987; Light & Singh, 1987).

In contrast, indirect tests of memory, which do not ask people to report on an event but require them to engage in some task that indirectly reflects the occurrence of that event, often do not reveal effects of aging. The elderly do not show significant memory deficits whether they perform perceptual identification tasks (Light & Singh, 1987), generate category exemplars in response to category names (Light & Albertson, 1989), or complete word stems or fragments (Light & Singh, 1987; Light, Singh, & Capps, 1986). These tasks are indirect because subjects are initially presented with items that serve as potential responses for the subsequent task but are not asked to think back to the earlier presentation during task performance.

The dissociation in performance on indirect and direct tests shown by the elderly has been interpreted as evidence that indirect tests reflect a form of memory or processing that aging spares (e.g., Howard, 1983; Light & Singh, 1987). Although discussed in the aging literature under a variety of terms such as priming (Rose, Yessavage, Hill, & Bower, 1986), procedural memory (Mitchell, 1989), and implicit memory (Light & Singh, 1987), the general notion is that performance on an indirect test does not entail deliberate recollection (Light & Albertson, 1989). Instead, the characteristics attributed to the processes underlying indirect test performance are similar to those ascribed to automaticity. Automatic processing has been described as a fast process that consumes no attentional capacity, is under the control of stimuli rather than intention, and occurs without awareness (e.g., Hasher & Zacks, 1979; Posner & Snyder, 1975; Schneider & Shiffrin, 1977). Dissociations between performance on indirect and direct tests have been described in terms of the contrast between automatic and consciously controlled processing (Jacoby, 1991; Klatzky, 1984; Logan, 1989).

Given this, it might be concluded that results from experiments in which indirect tests were used indicate that aging does not influence automatic uses of memory (cf. Hasher & Zacks, 1979). However, there are problems with that conclusion. Although the aged do not show significant deficits in performance relative to young subjects on the indirect tests mentioned earlier, there are consistent age trends favoring the young. Failures to find significant effects may have been due to insufficient power. Moreover, other studies have revealed significant age differences using indirect tests identical to those used to show age constancy (Chiarello & Hoyer, 1988; Davis et al., 1990; Howard, Shaw, & Heisy, 1986; Rose et al., 1986). Consequently, some researchers conclude that the form of memory underlying indirect test performance is not age invariant (Chiarello & Hoyer, 1988; Davis et al., 1990; Rose et al., 1986).

Conflicting interpretations of indirect test results stem from the assumption that each task taps a particular form of memory. Indirect tests are said to primarily reflect automatic or unconscious uses of memory, whereas direct tests primarily reflect strategic or consciously controlled uses of memory. However, there is good reason to believe that indirect tests are not process or factor pure (e.g., Jacoby, 1991; Richardson-Klavehn & Bjork, 1988; Schacter, 1987). Intentional, consciously controlled forms of processing may sometimes "contaminate" performance on indirect tests. Consequently, finding an age-related effect cannot be taken as evidence that aging produces a deficit in automatic uses of memory. The same problem exists when interpreting results on direct tests. Automatic forms of processing may contribute to direct test performance, lessening the impact of factors such as age and amnesia (e.g., Jacoby, 1991; Jacoby, in press).

In addition to the problem of contamination, isolating processes to a particular task may qualitatively alter the process being measured. Automatic and intentional uses of memory may be better examined in situations where both operate. Consider the problem of avoiding repeatedly telling a story to the same audience. Earlier tellings of the story likely have the effect...
of making the story come more readily to mind later. To avoid repetition, this automatic effect of memory must be successfully opposed by consciously recollecting an earlier telling of the story to the present audience: a use of memory that relies on recollection of source. Typically, memory for source is tested by directly asking people to report the source from which a given item or piece of information was gained. The elderly are impaired on these tasks (Cohen & Faulkner, 1989; Craik, Morris, & Morris, 1990; Hashtroudi, Johnson, & Chrosniak, 1989; McIntyre & Craik, 1987). However, there is an important difference between the requirements of a direct test of memory for source and the spontaneous monitoring of source that is required to avoid repetition (Dywan & Jacoby, 1990). Differences in performance on direct tests may underestimate age differences found in source monitoring (Dywan & Jacoby, 1990; Koriat, Ben-Zur, & Sheffer, 1988).

Avoiding repeating oneself provides a commonplace example of a situation that sets automatic and consciously controlled memory processes in opposition. In the following paragraphs, we describe the methodological advantages of arranging such a situation. Then we describe a process dissociation procedure that separates the effects of automatic and consciously controlled processes within a task. Rather than identifying different processes with different tasks, as is done by the direct-indirect test distinction, we examine different processes as they operate within a single task. Doing so allows us to study automatic influences as they operate in situations, such as source monitoring, where controlled processes are also in play. This type of examination could not be accomplished if one relied on an indirect test as a measure of automatic processing.

Advantages of Opposition

Indirect and direct tests that have been popular for investigating memory effects of aging typically constitute facilitation paradigms; that is, automatic influences of memory facilitate task performance just as would intentional, consciously controlled uses of memory. For fragment-completion performance, as an example, memory for the earlier reading of a word might automatically enhance performance, but intentional use of memory would produce the same result. Because the two types of processes produce effects in the same direction, age differences, using such tests, are difficult to interpret; effects might arise from a deficit in automatic processes, a deficit in consciously controlled processes, or both. Greater analytic power can be gained by arranging a situation such that automatic and consciously controlled processes produce opposite effects, that is, an interference paradigm. For interference paradigms, such as the task of avoiding repetitive recounting of a story, automatic uses of memory produce errors (e.g., repeated telling of the story) if left unopposed by consciously controlled processes (recollection for a previous telling). Deficits in consciously controlled processing, then, are reflected by an increase in errors.

In a series of false fame studies, automatic and consciously controlled uses of memory were set in opposition to examine the effects of aging and dividing attention. In the first phase of an experiment done by Dywan and Jacoby (1990), elderly and young adults read a list of nonfamous names, which were later mixed with famous and new nonfamous names and presented for fame judgments. Dywan and Jacoby correctly informed subjects that all of the names they read in the first list were nonfamous, so if they recognized a name on the fame test as one from the first list, they could be certain that the name was nonfamous. Earlier reading of a nonfamous name was expected to increase its familiarity, making it more likely that the name would later mistakenly be called famous. This automatic influence of memory, however, was opposed by conscious recollection that the name had appeared in the list (the source of the name's familiarity). Given this arrangement, any increase in the probability of mistakenly calling an old nonfamous name famous must result from an automatic influence of memory for its prior presentation, because conscious recollection would produce an opposite effect. Formally, old nonfamous names would mistakenly be called famous only if the name was familiar (F) but subjects did not recollect (R) the name as being presented earlier: F(1 — R). Elderly subjects were expected to show a deficit in recollection as compared with younger subjects and, consequently, to be more likely to mistakenly respond famous to old nonfamous names.

Results showed that elderly subjects were more likely to call old names famous than were younger subjects. Younger subjects used conscious recollection of the prior presentation of old nonfamous names to reject those names, whereas elderly subjects were less able to do so. Craik (e.g., Craik & Byrd, 1982) has argued that the memory effects of aging are similar to those produced by dividing attention of younger adults. In line with that possibility, Jacoby, Woloshyn, and Kelley (1989) showed that dividing attention either during the reading of nonfamous names or during the fame judgment test increased the probability of old nonfamous names mistakenly being called famous. Similar to aging, dividing attention makes it less likely that subjects will be able to recollect the prior presentation of old nonfamous names and, thereby, successfully oppose familiarity, an automatic influence of memory.

Process Dissociation Procedure

By placing familiarity and recollection in opposition, the false fame experiments showed that aging has the effect of reducing recollection, leaving effects of familiarity largely unopposed. However, from those results we cannot be certain that the elderly did not also show a deficit in automatic influences of memory (familiarity) as compared with younger subjects. The probability of calling an old name famous reflects a combination of automatic and intentional influences: F(1 — R). Unless recollection is fully eliminated (R = 0), the probability of calling old nonfamous names famous underestimates the influence of familiarity. To show that both aging and dividing attention influence recollection but have no influence on familiarity, we needed to separately estimate effects on the two types of processes, and we used the process dissociation procedure (Jacoby, 1991) to do so. Next, we illustrate that procedure by describing its application in Experiment 1.

In the fame experiment described earlier, subjects were told at test that earlier read names were nonfamous so as to place familiarity (F) and recollection (R) in opposition. We refer to that arrangement as an exclusion (Exc) test condition because
recollection served to exclude names that were earlier read. As described earlier, for an exclusion test, an old name will be mistakenly called famous only if it is sufficiently familiar (F) and not recollected as earlier presented \((1 - R)\). The probability of calling an old name famous on the exclusion test can then be represented as

\[
\text{Exc} = F(1 - R). \tag{1}
\]

In contrast, suppose we misinformed subjects that all of the earlier read names were actually obscure famous names. In this case (the inclusion [Inc] test condition), both recollection and familiarity would produce judgments of famous. That is, in contrast with the exclusion condition, recollection would serve to include earlier read names as famous. For an inclusion test, an old name could be judged famous either because it was recollected as being on the earlier read list \((R)\) or, although recollection failed, because the name was sufficiently familiar to be accepted as famous: \(F(1 - R)\). The probability of calling an old name famous on the inclusion test can then be represented as

\[
\text{Inc} = R + F(1 - R). \tag{2}
\]

Combining results from the inclusion and exclusion test conditions allows one to separately estimate the effects of intentional and automatic processes. Subtracting the probability of calling an old name famous on the exclusion test from that probability on the inclusion test provides an estimate of the probability of recollection:

\[
R = \text{Inc} - \text{Exc}. \tag{3}
\]

Given an estimate of recollection, an estimate of familiarity can be computed by means of simple algebra. One way of doing this is to divide the probability of calling an old name famous in the exclusion condition by the estimated probability of a failure in recollection \((1 - R)\):

\[
F = \text{Exc}/(1 - R). \tag{4}
\]

The probability of recollection can be best understood as a measure of intentional, consciously controlled processing defined in terms of selective responding. For the inclusion test, people are to select for old names, whereas for the exclusion test, people are to select against old names. Therefore, if the probability of recollection were 1.0, people would always call old names famous on the inclusion test and never call those names famous on the exclusion test. In contrast, if the probability of recollection were 0, people would be as likely to call an old name famous on an exclusion test as they would on an inclusion test.

We call this the process dissociation procedure because we are looking for factors that produce dissociations in their effects on the estimates of the different types of processes. It is important that we be able to find such dissociations. One of the strongest assumptions underlying the procedure is that automatic and intentional uses of memory are independent. If that assumption is valid, then we should be able to identify factors that have a large influence on one process but leave the other process unchanged. More complete descriptions of the concerns that led to the development of this procedure, its underlying assumptions, and its general utility are provided elsewhere (Jacoby, 1991; Jacoby, in press; Jacoby, Toth, & Yonelinas, 1993).

Because the process dissociation procedure allows us to separately estimate the contributions of automatic (familiarity) and consciously controlled (recollection) processes, we can separately examine the effects of aging on each process. In the first experiment, we examined the effects of both age and divided attention on familiarity and recollection using the fame judgment paradigm. The second experiment was similar, but we used a recognition memory task. On the basis of results from earlier experiments (e.g., Jacoby, 1993), we expected both dividing attention and aging to reduce the probability of recollection but have no influence on familiarity. That is, we expected an automatic influence of memory, familiarity, to be unaffected by aging and attention.

**Experiment 1**

In Experiment 1, we used a fame judgment task to examine the effects of aging and dividing attention. In Phase 1, a group of elderly and a group of young subjects devoted full attention to reading a list of nonfamous names. A second group of young subjects read the same list of names but under conditions of divided attention. For all three groups, an inclusion test was given in Phase 2 of the experiment, followed by an exclusion test in Phase 3. For the inclusion test, subjects were told that all old names appearing on that test were actually names of famous people, and so if they recollected reading a name in the earlier list, they should call the name famous. For the exclusion test, subjects were told that all old names appearing on that test were nonfamous, and so if they recollected reading a name in the earlier list, they could be certain that the name should be called nonfamous. Given results from these two test conditions, the equations described earlier were used to separately estimate the contributions of recollection and familiarity to fame judgments for each of the three groups of subjects.

It should be noted that one potential worry is that always giving the inclusion test before the exclusion test might produce a bias in the estimation procedure. An assumption underlying the procedure is that the probabilities of recollection and familiarity are the same for the inclusion and exclusion tests and that assumption may be violated because of forgetting across the tests. However, to examine the effects of aging and dividing attention it is not necessary that the estimates be totally unbiased so long as they are not differentially biased for the different conditions. We used the constant test order to simplify the design and because the results of earlier experiments (Jacoby, 1993) showed that order effects were not problematic.

**Method**

**Subjects.** The subjects were 54 young adults, ranging in age from 18 to 23 years \((M = 19.5)\), and 31 older adults, ranging in age from 66 to 90 years \((M = 73.8)\). The young subjects were enrolled in an introductory psychology course and took part for course credit. They had a mean of 12.9 years of education. Twenty-four of these subjects participated under a full-attention condition, whereas the remaining subjects participated in a divided-attention condition. Data from 6 subjects in this latter group were not used because those subjects were unable to perform at a level of 60% accuracy on the divided-attention task. Conse-
sequently, it is questionable whether their attention was truly divided.
The elderly adults were community-dwelling residents in the Hamilton, Ontario, Canada, area, who reported themselves to be in good health. They had a mean of 1.2 years of education. In addition, they had a mean score of 77% on the Mill Hill Vocabulary Scale (Raven, 1965). This score is comparable or better than that typically achieved by young adults. For example, in Experiment 2, the young adults had a mean score of 63%. Some of the elderly individuals came from a seniors’ group at a local church, whereas the rest came from the Seniors’ Centre at the Young Women’s Christian Association (YWCA). They offered their services voluntarily. Data for 7 elderly adults had to be discarded. These individuals were unable to follow instructions, as evidenced by their inability to show any discrimination between old and new test names, or to discriminate between famous and nonfamous names.

In summary, data were obtained from 48 young adults (24 individuals tested under a full-attention condition and 24 individuals tested under a divided-attention condition) and 24 elderly adults.

Materials. The stimuli were 126 nonfamous names and two sets of 120 famous names. The famous names were selected to be ones that subjects would likely recognize as famous without being able to identify what the named individual had done to achieve fame. This criterion was used to encourage subjects to base their fame judgments on the name’s familiarity rather than on its identifiability. Because names that were only familiar to the young subjects, such as Jack Benny or Betty Grable, were well-known to the elderly subjects, two sets of famous names were required. Each age group had their own set, developed from surveys conducted before the experiment.

The 126 nonfamous names were taken either from lists used in prior experiments (Dywan & Jacoby, 1990) or from the telephone book. These names matched the famous names on the following characteristics: gender as indicated by the first name, nationality of first and last names, the number of first and last names beginning with a given letter, and the length of the first and last names. Examples of nonfamous names are Sandra Baker and Wilson Love.

The nonfamous names were divided into two sets of 60. One set represented the “old” nonfamous names, presented at both study and test, whereas the second set was used as “new” nonfamous names, shown only at test. (The remaining six nonfamous names were used as filler items presented at the beginning of the study list.) Because there were two test phases, each set was further subdivided into two sets of 30: one for the inclusion test and the other for the exclusion test. Four formats were formed by rotating sets of names through conditions (old vs. new and inclusion vs. exclusion) such that, across formats, each name represented each combination of conditions. The famous names were presented only at test and were divided into two sets of 60, one set for each test. For both study and test, the presentation order of names was random with the restriction that not more than three names representing the same condition (e.g., study type or test type) could be presented successively.

Procedure. An Apple IIIE computer was interfaced with a monochrome green monitor to present the stimuli. The character size of the stimuli was approximately 5.7 x 6.6 mm. Names were presented in lowercase letters, with the initial letter of each first and last name capitalized, in the center of the screen. In the study phase, each name was presented for 2 s, and subjects were instructed to read the names aloud. They were told that their ability to pronounce the names quickly and accurately was of interest and that their pronunciation of the names was being recorded. In reality, no recording took place. Moreover, the subjects were given no indication that their memory for the names would be tested later.

The divided-attention subjects performed a listening task, previously used by Craik (1982), while reading names aloud. For that listening task, subjects monitored a tape-recorded list of digits to detect target sequences of three odd numbers in a row (e.g., 9, 3, 7). The digits were random with the exception that a minimum of one and a maximum of five numbers occurred between the end of one and the beginning of the next target sequence. The digits were presented at a rate of one digit every 1.5 s. Subjects pressed a key to indicate when they detected a target sequence. When subjects missed more than two consecutive sequences, they were prompted with the word “miss.”

Immediately after subjects read the list of nonfamous names, they were given an inclusion and then an exclusion fame judgment test. For those tests, subjects indicated that a name was famous by pressing one key or nonfamous by pressing a different key. Subjects were instructed to respond as quickly and accurately as possible. For the inclusion test, subjects were misinformed that all names in the test list that had been read in the earlier phase of the experiment were famous, and so if they recognized that a name had been read earlier, they should call it famous. For the exclusion test, subjects were told that only all earlier read names in the test list were names of nonfamous people, and so if they recognized that a name had been read earlier, they should call it nonfamous.

The significance level for all tests was set at p < .05. Tests revealing significant main effects are not reported when variables producing those main effects entered into significant interactions.

Results and Discussion

In the study phase, the young, divided-attention subjects missed a mean of 5.3 out of 33 target sequences (16%) in the listening task.

Although different sets of famous names were presented to the young and elderly adults at test, the probability of calling a famous name famous was comparable for the young, full-attention group (.77), the young, divided-attention group (.75), and the elderly group (.74). F(2, 69) = .189, MS_ε = 0.037. The probability of correctly calling a famous name famous between the inclusion (.76) and exclusion (.74) conditions also did not differ significantly, F(1, 69) = 3.75, MS_ε = 0.004, despite the use of different sets of names across the two tasks.

Examination of the probability of calling a new nonfamous name famous (Table 1) showed that the young, full-attention group was slightly less willing to call these names famous on the basis of familiarity than were subjects in the other two groups. Also, there is some indication that the elderly subjects used a higher criterion for familiarity-based judgments on the exclusion test than on the inclusion test. However, these differences were too small to be significant. In an analysis of the probability of calling a new nonfamous name famous, there

<table>
<thead>
<tr>
<th>Group</th>
<th>Inclusion names</th>
<th>Exclusion names</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Old</td>
<td>New</td>
</tr>
<tr>
<td>Young, full attention</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young, divided attention</td>
<td>.59</td>
<td>.22</td>
</tr>
<tr>
<td>Elderly adults</td>
<td>.58</td>
<td>.29</td>
</tr>
</tbody>
</table>

Table 1
Observed Probabilities of Calling Old and New Nonfamous Names Famous in Experiment 1
was no significant main effect of group, $F(2, 69) = 1.474$, $MS_X = 0.040$; test, $F(2, 69) = 1.037$, $MS_X = 0.010$; nor a significant Group $\times$ Test interaction, $F(2, 69) = 1.746$, $MS_X = 0.010$.

Estimates of discriminability for famous names were computed by subtracting the probability of calling a new nonfamous name famous from the probability of calling a famous name famous. Analysis of these difference scores revealed a significant Group $\times$ Test interaction, $F(2, 69) = 4.5$, $MS_X = 0.008$. A post hoc Newman-Keuls analysis suggests that the elderly adults showed less discriminability on the inclusion test than did the other two groups; however, there was no difference on the exclusion test. Reduced discriminability for the elderly may have stemmed from the use of different famous names between the age groups and across the tests. However, given that there were no group differences in the exclusion condition, it seems unlikely that differential discriminability of famous names had any effect on judgments of nonfamous names.

Of greater interest, analysis of the probability of calling old nonfamous names famous (see Table 1) revealed a significant Group $\times$ Test interaction, $F(2, 69) = 16.71$, $MS_X = 0.018$. The form of that interaction provides evidence that young, full-attention subjects held an advantage in recollection over subjects in the other two groups. For the inclusion test, recollection served to increase the probability of an old name being called famous. A post hoc Newman-Keuls comparison showed that the young, full-attention group was significantly more likely to call old names famous on the inclusion test than were the young, divided-attention group and the elderly adults. The latter two groups did not differ. For the exclusion test, recollection of a name as earlier read allowed subjects to be certain that the name was nonfamous and therefore decreased the probability of old names being called famous. The young, full-attention subjects called fewer old names famous on the exclusion test than did the young, divided-attention group or the elderly adults. The latter two groups again performed similarly.

More conclusive evidence of a difference in recollection between the three groups of subjects was gained by using the process dissociation procedure to separately examine effects on recollection and familiarity-based judgments. The probability of recollection was estimated as the difference between the estimated probability of calling a name famous on the basis of recollection versus familiarity. The estimated probabilities in Table 2 were computed using the means in Table 1 to allow the reader to verify use of the equations. However, for purposes of analyses, estimates were computed separately for each subject. The mean estimates of recollection computed for individual subjects were identical to the estimates computed from the group means. Estimates of familiarity were also similar (see Table 2). Analysis of the subject estimates revealed a significant difference between the groups in the probability of recollection, $F(2, 69) = 16.85$, $MS_X = 0.036$. According to post hoc Newman-Keuls comparisons, the young, full-attention group differed significantly from the other two groups, who showed a comparable level of performance. The results of those analyses show that both aging and divided attention produced large decrements in the probability of recollection, a consciously controlled form of processing.

In contrast with effects on recollection, the probability of calling an old name famous on the basis of its familiarity did not differ across the three groups, $F(2, 69) = 1.52$, $MS_X = 0.036$. This estimate of familiarity consists of automatic influences from prior presentation of the names at study and the baseline probability that names are familiar without prior exposure (the probability of calling a new name famous across the two tests). Presentation increased the familiarity of old names relative to the familiarity of new ones. The estimated probability of calling old names famous because of familiarity (Table 2) was significantly higher than the probability of calling new names famous (Table 1) for the young, full-attention group, $t(23) = 4.84$; the young, divided-attention group, $t(23) = 8.37$; and the elderly adults, $t(23) = 3.85$. That difference between old and new names reflects the influence of familiarity for earlier reading of the old names.

One can also examine differences in responding on the basis of familiarity by subtracting out the baseline probability of calling a name famous from the familiarity estimates. When we consider these values, the means for the young, full-attention (14), young, divided-attention (16), and elderly (13) groups are even more similar than when only the familiarity values for old names (Table 2) are examined. This is because, as mentioned earlier, the young, full-attention subjects called fewer new nonfamous names famous on the basis of familiarity than did subjects in the other two groups.

A vital assumption underlying the process dissociation procedure is the independent contribution of familiarity and recollection to memory performance. This assumption seems justified given that familiarity remained invariant with changes in age and attention, whereas recollection showed a sharp decline. However, to increase our confidence in this assumption, Pearson product-moment correlations between familiarity and recollection were calculated for each group. These correlations were not statistically significant for the young, divided attention group (28), young, full-attention group (29), nor the elderly adults (15), suggesting that the probabilities of basing a decision on recollection versus familiarity were independent.

In summary, the results of Experiment 1 showed that both aging and dividing attention produced large deficits in conscious recollection but left the use of familiarity as a basis for judgments unchanged. The effects on recollection are consistent with results from earlier experiments (Jacoby, Kelley, Brown, & Jacechko, 1989; Jacoby, 1993) and provide support for claims that aging and dividing attention have their effects by

Table 2

<table>
<thead>
<tr>
<th>Group</th>
<th>Recollection</th>
<th>Familiarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young, full attention</td>
<td>.60 (.60)</td>
<td>.33 (.31)</td>
</tr>
<tr>
<td>Young, divided attention</td>
<td>.34 (.34)</td>
<td>.38 (.39)</td>
</tr>
<tr>
<td>Elderly adults</td>
<td>.31 (.31)</td>
<td>.39 (.38)</td>
</tr>
</tbody>
</table>

Note: The mean estimates of recollection and familiarity calculated from individual scores are presented in parentheses.
limiting the possibility for consciously controlled processing. These results also confirm findings that suggest declines in recollection with age produce an increased false fame effect for names (Dywan & Jacoby, 1990) and faces (Bartlett, Leslie, Tubbs, & Fulton, 1989).

More important, perhaps, the results of Experiment 1 provide evidence that the use of familiarity as a basis for judgments remained invariant across differences in age and the manipulation of attention. The lack of effect of aging on familiarity-based judgments agrees with findings of age constancy on indirect tests of memory (for reviews, see Craik & Jennings, 1992; Light, 1991). However, because we used the process dissociation procedure, our conclusions are not based on the questionable assumption that tests are process pure.

Lastly, the influence of aging and divided attention on recollection and familiarity proved very similar. This finding corresponds to other results that show dividing attention in young adults produces effects in performance similar to those found with the elderly (e.g., Craik & Byrd, 1982). We further discuss the results of Experiment 1 after reporting Experiment 2.

Experiment 2

The results of Experiment 1 showed that dividing attention and aging both reduced the probability of recollection but left invariant the use of familiarity as a basis for fame judgments. In Experiment 2, we used a recognition-memory task to further examine the effects of aging. Dual-process theories of recognition memory (e.g., Atkinson & Juola, 1974; Jacoby & Dallas, 1981; Mandler, 1980) hold that judgments of familiarity and memory search or recollection serve as alternative bases for recognition-memory decisions. Just as for fame judgments, the use of familiarity for recognition-memory decisions may be an automatic influence of memory that is not affected by aging. That is, deficits in recognition-memory performance for elderly subjects, as compared with younger subjects (e.g., Light & Singh, 1987; Light, Singh, & Capps, 1986; White & Cunningham, 1982), may stem solely from a deficit in recollection, a consciously controlled use of memory. To examine that possibility, we used the process dissociation procedure to separately estimate the contributions of familiarity and recollection to recognition-memory judgments. Similar to results for fame judgments, we expected aging to have the effect of reducing recollection but to have no influence on the use of familiarity.

In Phase 1 of Experiment 2, a group of young subjects and a group of elderly subjects were presented with a list of words that were to be read intermixed with anagrams that were to be solved. In Phase 2, subjects in both groups heard a list of words that they were told to remember for a later test. An exclusion test then followed, and an inclusion test was given. In Phase 4 of the experiment. In both test phases, we used a two-alternative, forced-choice recognition test. One of the alternatives in each test was always an old word (one that was read or seen as an anagram in Phase 1 or heard in Phase 2), whereas the second alternative was always a new word that had not been presented earlier in the experiment. For the exclusion test, subjects were misinformed that each test pair included a word that had been presented. They were further told that they could make a correct choice by exclusion if they recollected a member of a pair as one that had been presented as a word to be read or as an anagram to be solved in Phase 1. Because words occurring in Phase 1 were never presented in Phase 2, subjects were told that if they recollected a word as occurring in Phase 1, they should pick the other word in the test as being the one that had been aurally presented. Given those exclusion instructions, subjects should select words that occurred in Phase 1 only if those words were familiar (F) but not recollected (1 - R) as earlier read or as solutions produced for anagrams.

For the inclusion test, subjects were informed that one member of each test pair was a new item and were instructed to pick the alternative that they recognized as occurring in either Phase 1 or Phase 2. For that test, a word that had been read or produced as a solution for an anagram in Phase 1 could be selected because subjects were able to recollect its prior occurrence (R) or because the word was sufficiently familiar (F) to be selected although recollection failed (1 - R). As in Experiment 1, the equations presented earlier were used to separately estimate the contributions of recollection and familiarity to recognition-memory performance.

The use of a forced-choice procedure was meant to eliminate any differences in criterion between elderly and young subjects. The recognition-memory task used in Experiment 2 also places heavier demands on recollection than does the fame judgment task used in Experiment 1. For the recognition-memory task, it is necessary to recollect whether a word was presented aurally or presented visually to be read or solved as an anagram. In contrast, for the fame judgment task, recollecting that a name was presented earlier, without reference to the details of that presentation, was sufficient. The more difficult recollection demanded for recognition memory might reveal even larger effects of aging on recollection than were observed in Experiment 1. One further difference between the two experiments is the order of testing. In Experiment 1, the exclusion test followed the inclusion test; however, in Experiment 2, the exclusion phase was given before the inclusion phase. If the results of Experiment 2 show the same pattern of performance as Experiment 1, we can be more certain that test order does not violate our assumptions about recollection and familiarity.

In other experiments, we have used procedures such as those used in Experiment 2 to separately examine effects on the different bases for recognition memory. Jacoby (1993) found that dividing attention during study reduced the probability of recollection but left invariant the use of familiarity as a basis for recognition-memory judgments. Jacoby (1991) also found effects of whether a word was read or produced as a solution for an anagram. Unlike the effects of dividing attention, reading words versus solving anagrams influenced both recollection and familiarity; solving anagrams produced a higher probability of recollection and greater familiarity of the solution words than did earlier reading those words. This issue will be further discussed after reporting the results of Experiment 2.

Method

Subjects. Two groups of subjects, a group of 20 elderly adults, ranging in age from 64 to 77 years (M = 70.2), and a group of 16 young
adults, ranging in age from 17 to 20 years (M = 18.9), participated in the experiment. The young adults were enrolled in an introductory psychology course and participated for course credit. They had a mean of 12.8 years of education and a mean score of 63% on the Mill Hill Vocabulary Scale (Raven, 1965). The elderly adults were community-dwelling residents in the Hamilton, Ontario, Canada, area, with self-reported good health. They had a mean of 11.6 years of education and a mean score of 70% on the Mill Hill Vocabulary Scale. These subjects came from the Seniors’ Centre at the YWCA and volunteered their help. Data for 2 of these subjects were lost because of problems with the computer program used for testing, whereas data for 2 other subjects could not be included because the subjects were unable to follow instructions on the exclusion test. In summary, data were obtained from 16 young and 16 elderly adults.

Materials. A pool of 204 five-letter words were chosen as stimuli. Sixty of those words were divided into two sets of 30 and used to construct study lists for Phase 1. One set of words was presented as anagrams to be solved; the second set was presented as words in their normal form to be read. These items were randomly intermixed for presentation. To construct the anagrams, words were presented with the second and fourth letters underlined and in their proper places, with the remaining letters randomly rearranged (e.g., “imse” for “smile”). Constraining the order of letters made the anagrams easier to solve and gave each anagram only one solution. Eight filler items (4 anagrams and 4 words in normal form) were shown at the beginning of the study list, creating a 68-item study list. A second set of 68 words was chosen from the stimulus pool for Phase 2 of the experiment. These words were presented aurally. The first 8 words in that list served as filler items. The remaining 68 words were used as distractors for the recognition tests in Phases 3 and 4. Each word was matched with one of the visual or auditory items from the study phase according to word frequency counts determined by Kucera and Francis (1967).

Two recognition test lists, one for each test phase, were constructed. Each list consisted of 30 visual items (15 from the anagram list and 15 from the to-be-read list) and 30 heard items. Each study item was paired with a distractor. In addition, each list began with eight of the filler items (two anagram items, two read items, and four heard items). These were used for practice trials to ensure that subjects understood the instructions.

Four formats were formed by rotating the Phase 1 items through each presentation and test condition (read vs. anagram and inclusion vs. exclusion). The Phase 2 items were also rotated through each test condition. The filler items, however, remained constant across the study and test conditions. The presentation order of words for both the study and test conditions was random with the constraint that no more than three items representing the same condition (e.g., type of study or test item) could be presented successively.

Procedure. An Apple IIE computer was interfaced with a monochrome green monitor to present the stimuli. The character size of the stimuli was approximately 5.7 x 6.6 mm. Stimuli were presented in lowercase letters in the center of the screen. In Phase 1, subjects were asked to solve anagrams and read words. They were told that the anagrams would be presented with the second and fourth letters underlined and that those letters were in their correct positions; therefore, only the remaining letters had to be rearranged. If subjects correctly solved the anagram, the experimenter pressed a key, initiating presentation of the next item; however, if subjects made an error, they were told to keep trying. A maximum of 30 seconds was allowed before subjects were told the solution and were asked to ensure that it was correct. The subjects were also informed that when words were presented in their normal form, their task was to read those words aloud as quickly as possible. The subjects believed that we were interested in their ability to solve anagrams and read words; they were not told that a recognition test would follow. In Phase 2, the remaining 68 study words were presented by an audio cassette recorder. Words were presented one at a time at a rate of one word every 2 s, and the subjects were asked to repeat each word aloud and try to remember it for a recognition-memory test.

In Phases 3 and 4, which immediately followed Phase 2, a two-alternative, forced-choice recognition test was presented on the computer. One of the two words was always an old word (i.e., one subjects had read, seen as an anagram, or heard), whereas the second word in the pair was always a new word. In the exclusion test (Phase 3), subjects were misinformed that each pair always contained a heard word, and they were instructed to choose the word they had heard by indicating whether it was on the left or right of the pair by pressing the corresponding response key. They were also told that if they recognized a word as one they had read or solved as an anagram, they should avoid that word and select the opposite one. For the inclusion test (Phase 4), subjects again had to complete a forced-choice recognition test. This time they were correctly informed that each pair consisted of an old and new item, and they were instructed to select any word they recognized as old, whether they had heard, read, or seen it as an anagram. Words that had been presented visually were now to be included. Both tests began with eight practice trials combined with feedback to make certain that the instructions were clear. Subjects were asked to make their decisions as quickly as possible. Following the final test phase, they were given the Mill Hill Vocabulary Scale (Raven, 1965).

The significance level for all tests was set at p < .05. Tests revealing significant main effects are not reported when variables producing those main effects entered into significant interactions.

Results and Discussion

The elderly adults solved 75% of the anagrams presented in Phase 1, whereas the young adults solved 86%. Although, this difference was statistically significant, F(1, 30) = 4.66, MS = 0.026, we would not expect it to have a great impact on the recognition-memory results. Both groups solved the majority of the anagrams, and when subjects could not solve an anagram, the solution was provided.

The probability of choosing an old word for each study and test condition is shown in Table 3. The probabilities of choosing read–anagram words are of primary interest. For those words, there was a significant interaction of study and test condition, F(1, 30) = 83.11, MS = 0.015, as well as a significant Group × Test interaction, F(1, 30) = 7.99, MS = 0.008. Those interactions show that young subjects held an advantage in recollection over elderly subjects and that the prior presentation of words as anagrams was more likely to be recollected than was that of words that were read.

For the inclusion test, recollection would serve to increase the probabilities of read and anagram words being selected. Post hoc Newman-Keuls tests showed that for that test, young

Table 3 Observed Probabilities of Choosing Old Words Across Study and Test Conditions in Experiment 2

<table>
<thead>
<tr>
<th>Group</th>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Read</td>
<td>Anagram</td>
</tr>
<tr>
<td>Young</td>
<td>.70</td>
<td>.87</td>
</tr>
<tr>
<td>Elderly</td>
<td>.60</td>
<td>.80</td>
</tr>
</tbody>
</table>
subjects were more likely to select study words than were elderly subjects, and words presented earlier as anagrams were more likely to be selected than were words that were read earlier for both groups. For the exclusion test, recollection served to exclude words that were read earlier and words that were presented earlier as anagrams. For that test, post hoc Newman-Keuls tests showed that younger subjects were less likely to select study words across the two study conditions than were elderly subjects, and words that were read earlier were more likely to be selected than were words that were presented earlier as anagrams for both groups. In addition, there was a significant Group (young vs. elderly) × Study Item (read vs. anagram) interaction, \( F(1, 30) = 4.93, M_{S} = 0.008 \). However, that interaction is not readily interpretable because it collapses across the inclusion and exclusion test conditions, which make very different demands on subjects.

We gained further evidence of differences in recollection by using the equations presented earlier to separately estimate the contributions of recollection and familiarity to recognition-memory decisions. The estimates presented in Table 4 are the values computed from the group means, with estimates calculated from individual scores presented in parentheses. The probability of recollection was higher for words presented earlier as anagrams than for words that were read earlier, \( F(1, 30) = 8.30, M_{S} = 0.030 \). In addition, the probability of recollection was significantly higher for young than for elderly adults, \( F(1, 30) = 86.77, M_{S} = 0.051 \). There was no significant Group × Study Item interaction, \( F(1, 30) = 1.11, M_{S} = 0.030 \).

To examine differences in familiarity, the individual score for 1 young subject had to be replaced with the group mean. That subject performed perfectly on the inclusion test, choosing all the read-anagram words, and made no errors in exclusion. Consequently, recollection equaled 1.0, and familiarity was undefined because its computations involved division by zero (see Equation 4). In contrast with the difference in recollection, the estimated probability of selecting a word on the basis of its familiarity was identical (.62) for young and elderly subjects. Neither the main effect of read versus anagram nor the Read Versus Anagram × Age interaction approached significance in the analysis of effects on familiarity, \( F(1, 30) = 1.54, M_{S} = 0.024 \), and \( F(1, 30) = 1.79, M_{S} = 0.19 \), respectively. The probability of selecting words from Phase 1 on the basis of their familiarity was significantly above chance (.50) for both the elderly, \( t(15) = 4.8 \), and the young adults, \( t(15) = 4.8 \).

The results of Experiment 2 replicated those of Experiment 1 in showing that the effects of aging produced a large deficit in the probability of recollection, but left invariant the use of familiarity as a basis for judgments. The advantage in recollection of words presented earlier as anagrams over words that were read earlier was expected (Jacoby, 1991). However, Jacoby also found that words presented as anagrams held an advantage in familiarity over read words, whereas we did not find an effect on familiarity. Jacoby argued that the effect of read versus anagram on familiarity was important because the effect is opposite to a prediction made by theories that hold that familiarity primarily relies on the match in perceptual characteristics between the study and test versions of an item (e.g., Jacoby & Dallas, 1981; Mandler, 1980). Because words were presented in their normal form to be read on the recognition test, words that were read earlier should have held an advantage in familiarity over words that were presented earlier as anagrams.

Jacoby (1991) interpreted his finding as evidence that familiarity is not totally reliant on the match in perceptual characteristics between the study and test versions. It is unclear why we were unable to replicate his result. Perhaps it is important that we used a forced-choice test, whereas he used a yes-no test of recognition memory. Regardless, even no difference in familiarity between anagram and read words would be damaging to theories that emphasize the importance of perceptual similarity for recognition. In the General Discussion section, we further consider factors that are important for familiarity and discuss differences in automatic influences revealed by tests of recognition memory, perceptual identification, and stem-completion tests. A point we make there is that automaticity is task dependent.

Another analysis examined recognition differences for words that had been presented aurally in Phase 2 (Table 3). There was an age difference in the probability of choosing those earlier heard words, \( F(1, 30) = 14.66, M_{S} = 0.010 \). The young adults were more likely to select earlier heard words than were elderly adults. That result is understandable as further evidence that the elderly subjects suffered a deficit in their ability to recollect the prior presentation of a word as compared with younger subjects. Because subjects were instructed to choose earlier heard words for both tests, one would not expect any effect of type of test on the probability of those words being selected. However, subjects were more likely to select earlier heard words on the inclusion test as compared with the exclusion test, \( F(1, 30) = 5.52, M_{S} = 0.008 \).

The effect of type of test on the probability of selecting earlier heard words might be taken as evidence for false recollection. On the exclusion test, subjects may falsely recollect an item they have heard as an item they have read or seen as an anagram and, consequently, exclude that item rather than select

<table>
<thead>
<tr>
<th>Group</th>
<th>Recollection</th>
<th>Familiarity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Read</td>
<td>Anagram</td>
</tr>
<tr>
<td>Young</td>
<td>.19 (.19)</td>
<td>.64 (.64)</td>
</tr>
<tr>
<td>Elderly</td>
<td>.07 (.07)</td>
<td>.43 (.43)</td>
</tr>
</tbody>
</table>

Note. The mean estimates of recollection and familiarity calculated from individual scores are presented in parentheses.
it. For the inclusion test, however, such errors in recollection would not reduce the probability of choosing heard items because subjects were to select all study items. The effect of test type for earlier heard words was the same for elderly and young subjects (F < 1 for the interaction), so any bias in the estimation procedure produced by false recollection was the same for both groups. Because our conclusions concern differences between the elderly and young rather than absolute values of estimated probabilities, the absence of a differential effect is all that is required.

Similar to Experiment 1, Pearson product-moment correlation coefficients measuring the relationship between recollection and familiarity were computed. Once again, there were small, statistically nonsignificant correlations between recollection and familiarity for the young adults in both the read (.04) and the anagram (.03) conditions. The elderly adults also produced a nonsignificant correlation (.36) in the read condition; however, the correlation in the anagram condition (.50) proved to be significant.

As discussed earlier, solving anagrams at study can influence both familiarity and recollection (Jacoby, 1991). Therefore, it is possible that the relationship between familiarity and recollection for anagram items may be special, particularly for elderly adults (cf. Verfaellie & Treadwell, 1993). However, across our two experiments, the correlations between familiarity and recollection range from -.29 to .50, with the variation not appearing to be systematic. Nonetheless, some caution about our independence assumption may be warranted in this case.

In summary, despite the different testing procedures used in Experiments 1 and 2, there is a striking similarity in the results; both experiments show the same pattern of performance with age. This is especially important given the different test order used in each experiment, which confirms other results, suggesting that a fixed test order is not problematic for the process dissociation procedure (Jacoby, 1993).

General Discussion

The results of our two experiments agree in showing that the effects of aging produce a substantial decrement in the probability of recollection but leave invariant the use of familiarity as a basis for judgments. The lack of an effect of aging on familiarity-based judgments is consistent with the conclusion drawn from performance on indirect tests that aging spares some functions of memory (Howard, 1983; Howard, 1988; Light & Albertson, 1989; Light & Singh, 1987; Light et al., 1986; Mitchell, 1989). Moreover, our results suggest that discrepant results found with the elderly on indirect tests stem from age differences in consciously controlled memory processes that are influencing performance on those tasks. Tests cannot be considered process pure, and measuring the effects of age on a specific process by performance on a single test does not always produce valid results.

Our conclusion that aging spares automatic processes but produces a deficit in consciously controlled processes is in agreement with arguments made by Hasher and Zacks (1979). However, by using the process dissociation procedure, we have redefined the terms consciously controlled and automatic. Jacoby, Ste-Marie, and Toth (1993) have argued that traditional definitions of automaticity are based on the assumption that tests are process pure, just as has been the interpretation of performance on indirect tests of memory Hasher and Zacks (1979), for example, argued that if memory for a particular attribute of an event is uninfluenced by instructions to remember, then processing of that attribute can be considered automatic. The interpretation of memory effects that are uninfluenced by instructions to remember suffers problems similar to those for interpretation of performance on an indirect test (cf. Begg, Maxwell, Mitterer, & Harris, 1986). Instead, we define automaticity in relation to a measure of consciously controlled processing. If a person is as likely to engage in a given act when trying not to as when trying to, then the person has no control. This intuitively appealing definition of control serves as the foundation for our measure of consciously controlled processing. In contrast, automatic processes do not support selective responding but, rather, produce the same effect regardless of one's intentions.

Our definition of automatic processing provides an integration for different historical phases in theorizing about the nature of the memory deficit suffered by the aged. In the 1970s, it was commonly held that the elderly were less resistant to effects of interference than were younger adults. During that phase, experiments presenting multiple lists to examine age differences in retroactive and proactive interference were common (for a review, see Winocur, 1982). The fragment-completion tasks, now popular as indirect tests of memory, were originally used by Warrington and Weiskrantz to restrict interference and, thereby, show savings in the performance of amnesics (see Weiskrantz & Warrington, 1975, for a description of that early work). The switch from the presentation of multiple lists to the presentation of fragments as retrieval cues is a change from an interference paradigm to a facilitation paradigm. Consequently, recent emphasis on preserved memory revealed by indirect tests can be seen as the "other side of the coin" of the greater interference effects suffered by aged subjects. In both cases, it is a deficit in consciously controlled processing in combination with relatively preserved automatic processes that produces the age-related effects. The procedure we have used here combines a facilitation paradigm (inclusion test) with an interference paradigm (exclusion test) to separately examine the effects of aging on the two types of processes.

Although we have redefined "automaticity" and "consciously controlled," by continuing to use those terms we mean to relate our work to current theorizing in the attention literature. One theme in that literature is that automatic processes are task and context dependent rather than totally stimulus driven (e.g., Logan, 1988; Neumann, 1984). Differences in automatic effects across task contexts can be seen by comparing automatic influences in stem-completion and perceptual-identification tasks with those in recognition-memory performance. For stem completion (Jacoby, Toth, & Yonelinas, 1993) and perceptual identification (Allen & Jacoby, 1990), reading words produces a larger automatic influence than does producing a word as a solution for an anagram, whereas either the opposite is true (Jacoby, 1991) or there is no difference between the two conditions (Experiment 2) in recognition memory. Because of differences in retrieval cues and task demands, automatic influences on recognition-memory judgments are less reliant on perceptual similar-
ity than are automatic influences on stem-completion and perceptual-identification tasks. Jacoby, Ste-Marie, & Toth (1993) have provided a more extensive discussion of the relativity of automaticity.

We hope to further refine the process dissociation procedure to produce standardized tests that can diagnose deficits in automatic and consciously controlled processes across a variety of task domains. To achieve that goal, a great deal more work is needed not only to refine procedures but also to explore differences in automatic and consciously controlled processes across tasks. We also believe that our approach points in new directions for the development of memory remediation programs. Typically, efforts to improve memory in the aged have adopted elaborate encoding schemes, using techniques such as pegword mnemonics (Wood & Pratt, 1987), method of loci (Kliegl, Smith, & Baltes, 1989; Robertson-Tchabo, Hausman, & Arendt, 1976), and face–name encoding (Yesavage, 1983; Yesavage, Rose, & Bower, 1983). Although these methods improve performance, the effects are usually task specific and shortlived (Scogin & Bienias, 1988; Wood & Pratt, 1987). We plan to train the elderly to improve their retrieval skills rather than their encoding skills. The process dissociation procedure provides a measure of conscious control that can be used as a target for training. Perhaps deficits in conscious control of memory are partially reversible by training just as are deficits in motor control (e.g., Duncan & Badke, 1987). Experiments to examine this possibility are in progress.

In conclusion, we used the process dissociation procedure to examine the effects of aging on automatic and consciously controlled memory processes as they operate within a single task. We found that age affects consciously controlled processes but leaves automatic memory processing intact. This pattern of results is the same as that generally reached by comparing performance on direct and indirect tests. A difficulty for the direct–indirect test distinction that we have circumvented, however, is the need to rely on the assumption that tasks are process pure. Rather than identifying processes with a particular type of task, the process dissociation procedure separates processes within a task. One further advantage that we have gained through this approach is the ability to examine the interplay of automatic and consciously controlled processes. We hope these advantages will be helpful in designing diagnostic memory tests and memory training procedures for the elderly.

References
Jacoby, L. L., Kelley, C. M., Brown, J., & Jaseckko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influ-

Received February 13, 1992
Revision received January 11, 1993
Accepted January 29, 1993