

## What underlies the deficit in reported recollective experience in old age?

TIMOTHY J. PERFECT and ZUBEIDA R. R. DASGUPTA  
*University of Bristol, Bristol, England*

In recognition memory, older adults report fewer occasions on which recognition is accompanied by recollection of the original encoding context. This study asks why this occurs. Two main hypotheses were tested: (1) Older adults fail to encode the items sufficiently when first they experience them. (2) Older adults have a retrieval deficit that prevents efficient reintegration of target and context. In addition, the hypothesis that frontal lobe integrity is essential for recollective experience was examined. Twenty older (mean age 70.7 years) and 20 younger (mean age 22.9 years) adults were asked to study a list of items and to verbalize the strategies they were using to remember. A further 20 older adults (mean age 70.0 years) were tested without the think-aloud protocol. Subsequently, subjects completed a battery of psychometric tests before completing a recognition test. As expected, older adults showed less recollective experience. They differed from the young in how they encoded the material, and once this difference was accounted for, no age differences in recollective experience remained. There was little evidence to support the hypothesis that frontal lobe integrity plays a role in the reduction of recollective experience with age.

This article is concerned with experiential changes in recognition memory that occur with increased age. It is widely acknowledged that increased age is accompanied by a quantitative reduction in the ability to remember past episodes (Salthouse, 1991), but the theoretical underpinning for this deficit is less well understood (Light, 1991). However, rather than focusing on age-related quantitative decline in memory, the present focus is for two reasons on the experience of memory. First, there is an intriguing pattern of findings regarding recollective experience in older adults that is worthy of investigation in its own right, and second, we believe that understanding the experiential changes in memory with old age will bring greater understanding of the quantitative changes that occur in later life. This belief echoes Koriat and Goldsmith's (1996) recent discussion of the correspondence and storehouse metaphors of memory, with the current focus on the experiential aspects (correspondence metaphor) rather than the amount of memory attained (storehouse metaphor).

To date, three studies have examined recollective experience in older adults using the Recollect/Know (R/K) procedure introduced by Tulving (1985). This procedure

involves a standard recognition test, followed by a decision by the subjects for each item they recognize, as to whether they remember or recollect the previous experience with the item (an R response), or merely find the item familiar, or know it was there, without recollecting the previous experience (a K response). The mean proportions of R and K responses from the studies by Parkin and Walter (1992), Mäntylä (1993), and Perfect, Williams, and Anderton-Brown (1995) are summarized in Table 1 to allow direct comparison.

Parkin and Walter (1992) and Perfect et al.'s (1995) Experiment 1 studied recollective experience for words, using instructions that did not specify how subjects should encode the material. The patterns of findings in these two studies are highly similar. For the younger subjects, around two thirds of recognition responses were accompanied by self-reports of recollective experience, with the remainder being found familiar only. The reverse was true for the older adults, with the majority of their recognition being based on familiarity only. Thus in these studies, old age was associated with a decrease in recollective experience and a corresponding increase in use of familiar-only responses.

Mäntylä (1993) reported a different pattern in his study of older adults. In this study, older and younger adults provided associates to test words at encoding, and these self-generated associates were later used as cues in a cued-recall paradigm. Subjects were asked to provide the original word to their cue and to make an R/K response to the item they had generated. What Mäntylä found that older adults showed less recollective experience to the items that they had generated than did the younger adults, but that the frequency of familiar-only responses did not differ between the age groups. Mäntylä attributed the re-

---

This work was supported in part by a grant from the Nuffield Foundation to T.J.P. The authors are grateful to Dan Wright for the advice given concerning the analysis of the data reported here, and to two anonymous reviewers who provided comments on an earlier draft. An earlier version of this paper was presented at an Economic and Social Research Council (E.S.R.C.) Workshop on Life Span Changes in Memory, in Worcester, U.K., May 1996. Correspondence should be addressed to T. J. Perfect, Department of Psychology, University of Bristol, 8, Woodland Road, Bristol BS8 1TN, England (e-mail: pstjp@ssa.bristol.ac.uk).

**Table 1**  
**Reported Proportion of Items Associated With**  
**Recollective Experience (R) and Familiarity in the Absence of**  
**Recollective Experience (K) in Previously Published Studies**

| Study   | Recollective Experience |     | Familiar-Only Responses |     |
|---|-------------------------|-----|-------------------------|-----|
|   | Rhit                    | Rfp | Khit                    | Kfp |
| Parkin & Walter (1992) Experiment 1   |                         |     |                         |     |
| Young   | .52                     | .01 | .25                     | .04 |
| Old   | .20                     | .01 | .46                     | .09 |
| Mäntylä (1993) Experiment 1   |                         |     |                         |     |
| Young   | .54                     | .05 | .11                     | .02 |
| Old   | .26                     | .05 | .10                     | .03 |
| Perfect et al. (1995) Experiment 1  |                         |     |                         |     |
| Young   | .53                     | .05 | .23                     | .05 |
| Old   | .18                     | .07 | .53                     | .03 |
| Perfect et al. (1995) Experiment 2A<br>(shallow and deep conditions combined) |                         |     |                         |     |
| Young   | .54                     | .01 | .33                     | .03 |
| Old   | .53                     | .03 | .14                     | .03 |

Note—fp, false positive.

duction in recollective experience to the more stereotypic cues that the older adults generated.

Perfect et al.'s (1995) Experiment 2A produced yet a different pattern. In this study, using the traditional depth-of-processing instructions to encode either count vowels or form an image, there was no deficit in recollective experience for older adults in comparison with the younger adults. It was this finding, compared with the marked recollective experience deficit in older adults in Experiment 1, that led Perfect et al. to conclude: "Without specific instructions older adults report high rates of recognition without recollective experience . . . because older adults do not spontaneously engage in elaborative encoding but rather spend their time engaging in additional physical analysis" (p. 185).

This argument was based in part on the work of Gardiner, Gawlik, and Richardson-Klavehn (1994), who, using a directed forgetting paradigm, showed that increased maintenance rehearsal leads to higher rates of familiar-only responses, whereas increased elaborative rehearsal leads to greater recollective experience. Nonetheless, as Perfect et al. (1995) acknowledged, the idea that older adults lack recollective experience because they do not spontaneously utilize deeper encoding strategies requires direct verification.<sup>1</sup> Thus the main focus of the present paper is the examination of the spontaneous encoding strategies of younger and older adults in order to test this hypothesis.

The hypothesis that older adults lack recollective experience because they fail to spontaneously encode material they encounter is related to, but not conceptually the same as, the account offered by Parkin and Walter (1992). Parkin and Walter attributed the reduction in recollective experience in older adults to an impairment of frontal lobe functioning, since in the oldest group of older adults there were correlations between measures of frontal lobe

functioning and amount of recollective experience reported. However, as Parkin and Walter acknowledged, an account based on the idea that recollective experience declines because of loss of frontal lobe functioning does not specify the mechanism by which this occurs. Impaired frontal lobe functioning may result in poorer encoding, a retrieval deficit, or both.

There are two problems in accepting the frontal hypothesis on the basis of the evidence put forward by Parkin and Walter (1992): (1) The correlations between frontal tests and recollective experience were found only in the oldest subjects (Experiment 1, mean age 80.0 years; Experiment 2, mean age 81.6 years), yet the reduction in recollective experience was also seen in a middle-old group (Experiment 2, mean age 67.7 years) that showed no frontal correlations with recollective experience. (2) There were no tests of nonfrontal skills to contrast with the frontal tests. Thus the reduction in recollective experience may have been related to general cognitive ability, rather than a specific decline in executive/frontal function. In order to examine this idea, subjects in the present study completed tests of frontal functioning (Wisconsin Card Sort Test, FAS fluency, cognitive estimates test), as well as tests of general ability not thought to load so heavily on executive functioning (vocabulary, mental rotation).

Our selection of frontal tasks was determined by three main constraints: We wished to provide a test of the frontal hypothesis that was consistent with previous work, we wished to test a range of putative frontal tests, but we did not wish to overtax the elderly volunteers with endless tests. Our decision was to use the two measures of frontal test performance originally used by Parkin and Walter (1992), namely the Wisconsin Card Sort Test and the FAS fluency test; the Wisconsin card sort taps the ability to inhibit dominant responses and avoid perseveration; FAS fluency taps the ability to access vocabulary rapidly and fluently. To this we added the cognitive estimates test, which involves making judgments about memory and has also been shown to be frontally sensitive (Shallice & Evans, 1978). All three tests tap frontal lobe involvement in "the manipulation and organization of information" (Shimamura, 1995, p. 808). However, they tap different aspects, so it would be a mistake to simply combine them into a single frontal measure. As Parkin and Lawrence (1994) have argued: "Human memory and frontal lobe functions are both complex entities and it seems unlikely, on intuitive grounds at least, that the relation between them should be unidimensional. More likely both memory and frontal function are multicomponential so that any given memory function might be associated with a particular subset of frontal functions. This idea has not, as yet, received any investigation" (p. 1524).

Our study therefore had two main purposes. First it aimed to determine whether there is evidence for frontal involvement in the deficit in recollective experience by contrasting the predictions from frontal and nonfrontal tests. With regard to the involvement of the frontal lobe

tests, our intention was to determine which tests of frontal function best pick up age-related variance in recollective experience. Second, the present study aimed to examine the mechanism of the effect by examining both the encoding strategy used by younger and older adults and the levels of recollective experience once encoding had been controlled. Our expectations were as follows: (1) Older adults would show less recollective experience. (2) Older adults would show less spontaneous elaboration and more use of maintenance rehearsal. (3) Controlling for measures of frontal lobe performance, but not nonfrontal lobe performance, would remove age differences in recollective experience. (4) Controlling for encoding strategy would remove age differences in recollective experience.

Since the technique we planned involved recording what the subjects had thought of at encoding, we were able to examine one further issue. At test, as well as asking subjects to report whether or not they had recollective experience for the original exposure to the item, we could ask subjects to tell us what they had originally thought of. This approach is not entirely new; Loftus and Kalman (1979) showed in a picture recognition paradigm that subjects who named features of pictures at study and test performed better (measured by  $d'$ ) than did those who named features at test only. However, this approach has not been used before in the context of recollective experience, and it will allow us to verify whether subjects who reported recollective experience were remembering the same thoughts that they originally vocalized. Of course, subjects may have not said out loud all the ideas that passed through their minds, and so our estimate provides us with only the minimal proportion of items for which we can verify that recollective experience is accurate.

An additional issue to be introduced is the materials we used to test the subjects. In the literature, there have been three demonstrations of increased levels of K responding. As discussed above, two of these concern the performance of older adults—Parkin and Walter (1992) and Perfect et al. (1995). The other demonstration of this pattern was by Gardiner and Java (1990), who showed that recognition memory for nonwords is accompanied by less recollective experience and more familiar-only responses than is recognition memory for words. Thus nonwords, which by their nature are hard to elaborate, result in the same pattern for younger and older adults with words. We therefore decided to examine older and younger and older adults on words and nonwords to see whether the age effect on words would be exaggerated when nonwords are used. We predicted that older adults would show a greater deficit in recollective experience for nonwords than for words.

There has been considerable recent debate as to how R and K responses should be interpreted, and it is important to deal with this issue prior to presentation of the data. The issues of debate are whether R and K responses represent evidence for one memory process (see, e.g., Donaldson, 1996) or two underlying processes (see, e.g., Rajaram, 1993), and what the relation between those pro-

cesses is. As Jones (1987) originally argued, with two processes, three models are theoretically possible: (1) An exclusivity model, in which there are two mutually exclusive processes, assumes that these processes are tapped directly by R and K responses. This was the view initially argued by Gardiner and Parkin (1990). (2) An independence model, in which there are two independent but not exclusive processes, assumes that R responses directly measure the process of recollection, but K responses measure only the output of the familiarity process, which is not recollected. This is the view proposed by Yonelinas and Jacoby (1995). (3) A redundancy model, in which recollection as a process is a subset of the familiarity process, posits that all items that are recognized are found familiar, but only a subset of these are recollected. Hence under this assumption R responses measure the proportion of familiar items that are recollected, whereas K responses measure the proportion of familiar items that are not recollected. The latter view is consistent with Donaldson's (1996) two-threshold signal-detection model, since items that have greater strength than the higher threshold (for recollection) automatically also have greater strength than the lower threshold (familiarity).

The present focus is on recollective experience. The models of exclusivity and independence are based on the assumption that the process of recollection can be directly tapped by the frequency, or proportion, of R responses. For that reason, we will use the raw R scores as measures of recollection. However, the redundancy model is based on a different assumption, and Donaldson (1996) has argued that this measure of recollection is best tapped by use of the nonparametric measure of memory strength,  $A'$ . For this reason, where appropriate we will cite the  $A'$  values in addition to the raw scores. We will draw attention to those occasions when the two models lead to different interpretations of our data. Because there is less agreement about the interpretation of K responses, and because our focus is on recollective experience, we will report the raw K responses only, before turning our attention to understanding the age-related changes in recollective experience.

There is a second important reason for retaining the use of the raw R scores as a measure of recollection. Our aim is to investigate the basis for R responses on an item-to-item basis. This is simply not possible if one makes transformations of the estimates of recollection on the basis of stochastic models. Thus in the later analyses we report, we explicitly use the assumption that the process of recollection is what is tapped directly by R responses.

## METHOD

### Subjects

Initially, 20 older and 20 younger adults were tested using the full procedure outlined below. A second group of 20 older adults, whom we did not tape-record at encoding, were later added to ensure that the tape-recording process had not altered the recognition memory performance we observed. No difference between the two older groups was observed, so the two groups were combined when appropriate. The subject details are given in Table 2.

**Table 2**  
**Mean Ages, Scores, and Standard Deviations on Mental Rotation Test, NART Vocabulary Test, Cognitive Estimates Test, FAS Fluency Test, and Percentage Perseverative Errors on Wisconsin Card Sort Test for Each Group**

| Group                  | Age      |           | NART     |           | ROT      |           | COG      |           | FAS      |           | Wiscp    |           |
|------------------------|----------|-----------|----------|-----------|----------|-----------|----------|-----------|----------|-----------|----------|-----------|
|                        | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> |
| Young ( <i>n</i> = 20) | 22.90    | 2.85      | 36.30    | 5.74      | 18.85    | 1.39      | 15.95    | 4.36      | 40.00    | 7.94      | 2.59     | 2.35      |
| Old 1 ( <i>n</i> = 20) | 70.65    | 4.84      | 37.45    | 8.06      | 17.25    | 1.83      | 16.90    | 5.41      | 34.15    | 8.08      | 6.58     | 5.43      |
| Old 2 ( <i>n</i> = 20) | 69.95    | 4.81      | 33.45    | 4.81      | 14.90    | 3.51      | 19.48    | 6.46      | 31.80    | 6.65      | 14.73    | 23.93     |
| 95% C.I.               | 45.4     |           | -4.54    |           | -3.91    |           | -0.50    |           | -11.34   |           | 2.23     |           |
| (old - young)          | 49.4     |           | 2.85     |           | -1.64    |           | 4.98     |           | -2.71    |           | 13.69    |           |

Note—Old 1, The older adults who were asked to use the think-aloud protocol and whose encoding strategies were recorded; Old 2, the older adults who did not use the think aloud protocol; 95% C.I., 95% confidence interval for the age difference (with older adults collapsed into a single group); NART, National Adult Reading Test; ROT, performance on the mental rotation test; COG, cognitive estimates test; FAS, FAS verbal fluency test; Wiscp, perseverative errors on the Wisconsin Card Sort Test.

### Materials

Forty words with a frequency range of 99–371 per million (Francis & Kučera, 1982) were selected. All were between 4 and 10 letters long. In addition, there were 40 pronounceable nonwords, generated by the experimenters (e.g., *pedon*, *frumstle*). Half the words and the nonwords were presented for study, mixed in random order. Order of presentation of words and nonwords was randomized, and the half that was presented first was counterbalanced within each age group.

### Test Battery

**The frontal tests.** Shallice and Evans's (1978) cognitive estimates questionnaire was administered to each subject and scored in the conventional manner. High scores on this test indicate poorer ability to make estimates. Items require estimates of distance, speed, weight, amount, and frequency (e.g., "How tall is the average British woman?" "What is the length of a £10 note?"). Frontal patients have been shown to be impaired at this task (Shallice & Burgess, 1991; Shallice & Evans, 1978).

In the FAS fluency measure (Benton, Hamsher, Varney, & Spreen, 1983), subjects are asked to write down as many words starting with the letter *F* as they could think of in 1 min. They are told that names of people, places, or multiple words with the same root are not allowed (e.g., *fun*, *funny*). This procedure is repeated for the letters *A* and *S*, and the total score of legal words produced is taken. This measure was used in the original Parkin and Walter (1992) study.

In the Wisconsin Card Sort Test (modified version, Nelson, 1976), subjects are required to discover the rule for sorting a set of cards on the basis of experimenter feedback. After six correct categorizations, the rule is changed, and subjects are required to change the basis of their categorization. There are three possible categorizations, and subjects have to complete six rule changes in all (twice through the three rules). Because of our interest in frontal lobe function, the measure taken was the number of perseverative errors made by subjects on the task. This measure was used in the original Parkin and Walter (1992) study.

### Nonfrontal Tests

In the National Adult Reading Test (NART; Nelson, 1982), subjects are required to read out loud a series of irregular words that decrease in frequency of use in the English language. The score obtained is the number of correct pronunciations obtained on the test.

In the rotation test, subjects are presented with pairs of line drawings of irregular geometric shapes drawn on flash cards. Subjects are asked whether one of the pair could be rotated so that it is identical to the other member of the pair. For half the pairs this is possible, whereas the remainder are mirror images of one another. Subjects are not allowed to physically manipulate the materials. In the present study, they were asked to work through the set of 20 pairs

as quickly and as accurately as possible. The score obtained was the number of correct classifications (match/mismatch) out of 20.

### Procedure

All subjects were tested individually in a small test room. Subjects were informed that the experiment concerned memory. Subjects were told that the experimenters were interested in *how* they remembered—that is, what they did in order to remember words and nonwords. They were told that they would be presented with a mixture of words and nonwords, one at a time on flash cards, for 5 sec per item, and asked to remember each item. Subjects were instructed to read each item out loud as they saw it and then say everything they were thinking of out loud in order to remember the item until the item was taken away. Each subject then conducted a number of practice trials until the experimenter was satisfied that he/she understood the instructions. The subjects' responses to the 40 experimental trials (20 words and 20 nonwords in random order) were then tape-recorded and later transcribed.

Following presentation of the to-be-remembered items, subjects completed the following tests in the given order: Wisconsin Card Sort Test, FAS fluency test, cognitive estimates test, and rotation test. Subjects then completed the recognition test for the words and nonwords before taking the NART vocabulary test.

At recognition, subjects were presented with a printed list of 80 items (40 words and 40 nonwords) in random order. Each item had the letters R/K/N printed next to it. Half the items had been presented previously, although the subjects were not informed of this. Subjects were asked to indicate for each item whether they remembered the prior occurrence of the item (by circling the R on the printed sheet), whether they knew the item had been presented earlier, but could not remember the actual occurrence (by circling the K), or whether they thought that it was new (by circling the N). Subjects had the R/K distinction carefully explained to them using the procedure outlined in previous studies (see Perfect, Mayes, Downes, & van Eijk, 1996). Once subjects had made a response for every item in the list, they were asked to go back to each item that they recollected and say what it was that they remembered about the original encoding. These reports were transcribed, compared with the original encodings, and then classified as being the same or different by the experimenter. A loose criterion was used, wherein any reference to a previously mentioned thought was counted as being the same, even if the phrasing was different.

The young subjects and the first group of older adults were tested using the procedure outlined above. A second group of older adults was subsequently tested without the requirement that they vocalize their thoughts during study of the items. They were merely asked to try to remember the items, and they were not tape-recorded at encoding. Consequently, they were not asked to explain their R responses. In all other respects the procedure was identical.

**RESULTS**

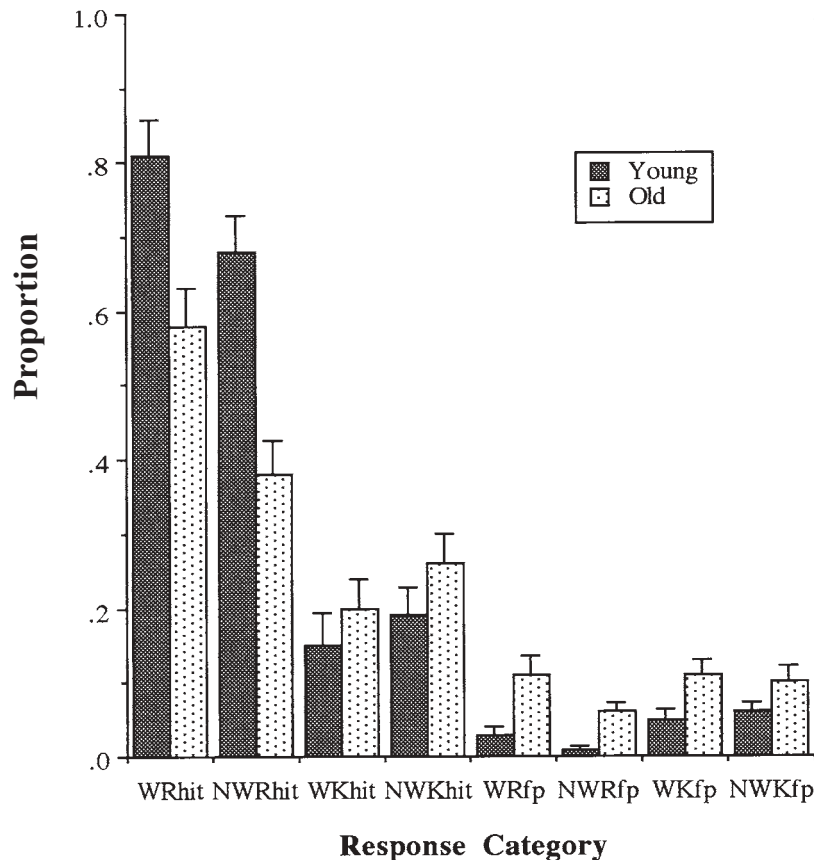
As outlined in the Method section, originally we tested 20 older and 20 younger adults, and we recorded their encoding strategies. Later we added a further 20 older adults whom we did not record. The original intention was to determine whether the act of recording altered the pattern of R and K responses observed. It did not, so in the following analyses the older recorded and non-recorded groups are combined to form a single group of older adults. Only the original 20 older adults are included in the later analyses, which examine the encoding strategies in detail.

The preliminary analyses concerned the age effects on the frontal and nonfrontal tests, shown in Table 2. As can be seen from the 95% confidence intervals shown at the foot of that table, the older adults did not differ from the younger adults on NART vocabulary, but were worse at mental rotation. This is entirely to be predicted, given the well-documented age pattern in fluid and crystallized intelligence (Horn & Catell, 1967). For the frontal tests, older adults were no worse overall on the cognitive estimates test, but were less fluent and made more perseverative errors on the Wisconsin Card Sort Test. The latter

two findings replicate previous findings (Parkin & Lawrence, 1994; Parkin & Walter, 1992).

Our main focus in this study was the level of recollective experience. However, in order to allow comparison with previous studies, our first analysis examined the raw rates of R and K responding. We therefore conducted  $2 \times 2 \times 2$  (age group  $\times$  item type [words vs. nonwords]  $\times$  response status [hits vs. false positives]) analyses of variance (ANOVAs) separately for R and K responses. The mean proportions of R and K responses are shown in Figure 1.

As Figure 1 shows, the older adults reported less recollective experience [ $F(1,58) = 6.41, MS_e = 31.04, p < .02$ ], there was less recollective experience for nonwords [ $F(1,58) = 18.85, MS_e = 11.11, p < .001$ ], and there was more recollective experience for targets than for distractors [ $F(1,58) = 303.5, MS_e = 22.22, p < .001$ ]. The analysis indicated an interaction between age group and response status [ $F(1,58) = 26.24, MS_e = 22.22, p < .001$ ]: Figure 1 shows that older adults had fewer recollective experiences to targets and more false recollective experiences. There was also an interaction between target type and response status; Figure 1 indicates that there were fewer correct recollective experiences for nonwords, but



**Figure 1.** Recognition performance for old and young groups, split into recognition with recollective experience (R) and recognition based on familiarity only (K). W, word; NW, nonword; fp, false positive.

no difference between words and nonwords for false recollective experience [ $F(1,58) = 12.85$ ,  $MS_e = 7.08$ ,  $p < .001$ ]. No other interaction was significant. For the frequency of items that subjects claimed to recognize with a sense of familiarity only, Figure 1 shows that the older group used the K response marginally more often than the young or group [ $F(1,58) = 2.88$ ,  $MS_e = 23.95$ ,  $p < .10$ ], but there was no effect of item type [ $F(1,58) = 1.40$ ,  $MS_e = 8.12$ ]. There were no significant interactions.

As outlined in the introduction, we decided to examine rate of recollection by analyzing both the raw rate of response and the nonparametric measure of discrimination,  $A'$ , as recommended by Donaldson (1996). Figure 2 shows the values of  $A'$  for both R responses alone and K responses alone. For the R responses, older adults were clearly impaired on  $A'$  for both words and nonwords [ $F(1,53) = 18.37$ ,  $MS_e = 0.02$ ,  $p < .001$ ], but there was no effect of age on  $A'$  for the K responses ( $F < 1$ ). Thus the pattern observed in the  $A'$  scores mirrors that seen in the raw data.

The remainder of this paper focuses entirely on the measures of recollective experience. Three regression analyses were conducted,<sup>2</sup> each with the same format. In each, the first step was to see whether age predicts variation in the dependent variable (recollective experience, encoding strategy, and recollective experience given en-

coding strategy, respectively). The second step was to see whether this age effect was eliminated by controlling for the individual difference measures (nonfrontal—NART, rotation test score; frontal—cognitive estimates, FAS fluency, and perseverative errors on the Wisconsin Card Sort Test). If controlling for frontal test scores eliminated the age effects, the frontal hypothesis would be supported. If controlling for frontal and nonfrontal test performance removes the age effect equally, this would constitute support for the generalized deficit model of aging.

The first analysis, presented in Table 3, shows the predicted variance in reported recollective experience, measured both by raw score and by  $A'$ . Age significantly predicts recollective experience, whether measured by hits and false positives separately, or by  $A'$ . With few exceptions, the age-related variance remained even after the measures of general cognitive ability and frontal test scores had been controlled for. The exceptions were that removing the mental rotation score removed the age effect for recollective experience for words, and false positives for nonwords, and the frontal measures of FAS fluency and perseverative errors on the Wisconsin Card Sort Test removed the age effect on hits for words.

The next analysis concerned the different spontaneous encoding strategies used. These were classified as elaborative or shallow,<sup>3</sup> for words and nonwords, with an ad-

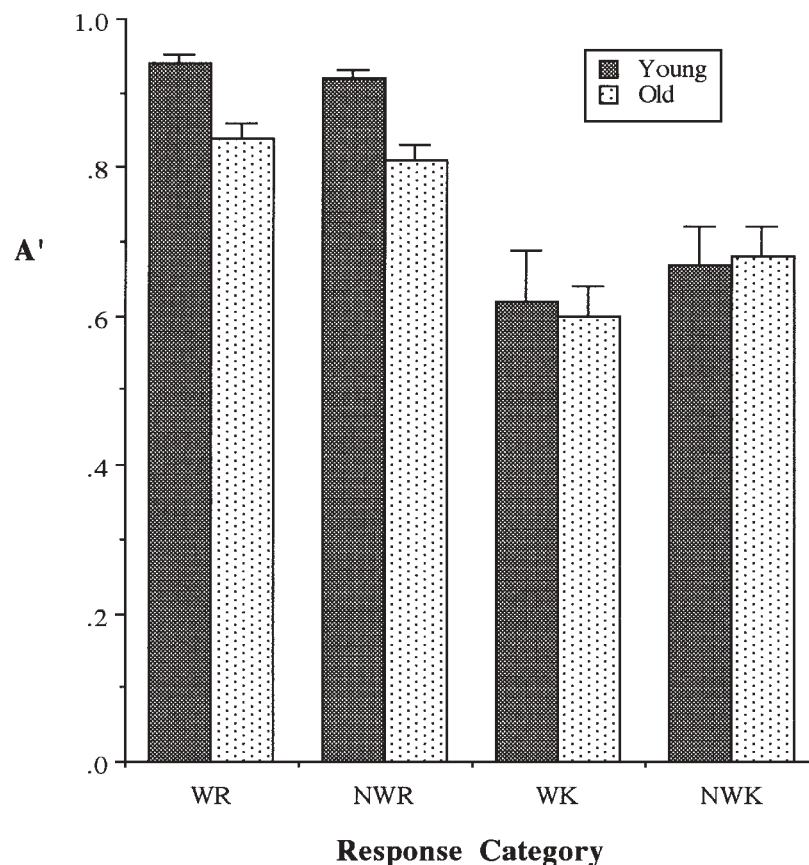


Figure 2. Recognition performance for old and young groups, measured by  $A'$  for R and K responses separately. W, word; NW, nonword; N, new.

**Table 3**  
**Variance in Frequency of Recollective Experience Explained by Psychometric Test Performance, Age, and Age After Psychometric Test Performance**

| Response Measure | Age Alone    | Age After Psychometric Predictors |              |              |              |              |
|------------------|--------------|-----------------------------------|--------------|--------------|--------------|--------------|
|                  |              | NART                              | ROT*         | COG          | FAS          | Wiscp*       |
| Words            |              |                                   |              |              |              |              |
| Hit              | <b>10.7†</b> | 7.7†                              | 2.0          | <b>8.6†</b>  | 5.4          | 4.8          |
| FP*              | <b>12.0‡</b> | <b>18.1§</b>                      | 5.4          | <b>14.1‡</b> | <b>10.1†</b> | <b>10.3†</b> |
| A'*              | <b>18.3§</b> | <b>18.7§</b>                      | 4.9          | <b>16.3‡</b> | <b>11.9†</b> | <b>9.7†</b>  |
| Nonwords         |              |                                   |              |              |              |              |
| Hit*             | <b>20.2§</b> | <b>18.7§</b>                      | <b>11.4†</b> | <b>15.3‡</b> | <b>10.5†</b> | <b>8.8†</b>  |
| FP*              | <b>10.3†</b> | <b>14.4‡</b>                      | 4.0          | <b>11.5†</b> | <b>13.9‡</b> | <b>12.3‡</b> |
| A'*              | <b>29.0§</b> | <b>33.2§</b>                      | <b>15.0‡</b> | <b>25.9‡</b> | <b>20.7§</b> | <b>17.8‡</b> |

Note—Bold type indicates a significant prediction. Hit, rate of correct identification of previous items that were given an R response; FP, rate of false positive responses to distractors that were given an R response; A', Nonparametric measure of discrimination for R responses. NART, National Adult Reading Test; ROT, performance on the mental rotation test; COG, cognitive estimates test; FAS, FAS verbal fluency test; Wiscp, perseverative errors on the Wisconsin Card Sort Test. \*Data were log-transformed to remove skewing prior to analysis. †*p* < .05; ‡*p* < .01; §*p* < .001.

ditional two categories of response: (1) A response was categorized as “blank” if the subject failed to say anything in response to the item, and (2) a response was categorized as “no strategy” if the subject explicitly stated that he/she could not think of a response to the item. The predictors of strategy use are shown in Table 4. There were two significant age effects on strategy use: Older adults used elaborative encoding less often for nonwords, and they used the “no strategy” response more often than did younger adults. However, as before, the age effect, with only one exception, was not removed by controlling for the measures of frontal test performance or the general cognitive measures. The exception was that the age effect on the use of deep encoding for nonwords was removed by controlling for FAS fluency. Thus it is clear that there are age effects in encoding, but it is less clear as to the genesis of these age differences, since they were not removed by either frontal or nonfrontal test scores.

The next analysis examined the likelihood of generating an R response when an elaborative encoding style had been used. Our initial intention was to analyze the results of both shallow and elaborative encoding, but the shallow encoding was not used frequently enough to support further analysis on the subset of responses that later elicited an R response. The data, shown in Table 5, clearly demonstrate that once encoding differences have been taken into account, there were no age effects on reported recollective experience for words or nonwords. Controlling for individual difference measures did not alter this pattern.

The final analysis, shown in the bottom row of Table 5, examined the proportion of R responses for which the reported basis of recollection was the same as that which had been given in the subjects' original think-aloud protocol at encoding. Here there was a clear age-related decline, with older adults producing verifiable R responses

on only 51% (*SD* = 27%) of the time, with the equivalent value for younger adults at 72% (*SD* = 22%). Here there was some support for the frontal hypothesis in that the age effect on ability to generate the same knowledge at recollection that had been reported at encoding was removed by controlling for both the FAS fluency measure and perseverative errors on the Wisconsin Card Sort Test. However, this analysis was conducted without regard to the nature of the encoding that the two groups had conducted. If the likelihood of using a deep encoding strategy is partialled out, the 21.5% of age-related variance in the *same* measure is reduced to a nonsignificant 4.3%. Thus this effect, which looks at first glance like a retrieval deficit, is in fact related to the nature of the encoding used.

**DISCUSSION**

The results showed the expected reduction in reported recollective experience with old age and the expected reduction in recollective experience for nonwords compared with words. However, the expected increases in familiar-only responses for older adults and for nonwords were not found. Thus the present aging data are similar to those of Mäntylä (1993), rather than those of Parkin and Walter (1992). This is somewhat unexpected, given the pattern of performance reported by Perfect et al. (1995) in their “no instructions” condition. Since no specific instructions on how to encode were given, an inflated rate of K responding was expected in the older adults. Also the lack of an interaction between age and item type was unexpected. Initially, we hypothesized that the think-aloud protocol was responsible for this, but the same pattern was observed for the second older group, which was not told to think aloud. It is possible that item effects caused this. Unlike the previous work with older adults, and with younger adults studying nonwords, the present study used a mixed list of words and nonwords together. Perhaps the mixing of items maintains attention or en-

**Table 4**  
**Percentage Variance Explained by Psychometric Tests, Age, and Age After Controlling for Psychometric Tests in Frequency of Use of Different Encoding Strategies**

| Encoding Strategy | Age Alone    | Age After Psychometric Predictors |              |              |              |              |
|-------------------|--------------|-----------------------------------|--------------|--------------|--------------|--------------|
|                   |              | NART                              | ROT*         | COG          | FAS          | Wiscp*       |
| Words             |              |                                   |              |              |              |              |
| Deep*             | 0.4          | 0.6                               | 0.2          | 0.1          | 0.2          | 0.1          |
| Shallow*          | 6.8          | 4.5                               | 7.3          | 4.8          | 6.1          | 7.0          |
| Nonwords          |              |                                   |              |              |              |              |
| Deep              | <b>25.5‡</b> | <b>17.4†</b>                      | <b>11.8†</b> | <b>17.9†</b> | 7.7          | <b>13.6†</b> |
| Shallow*          | 1.2          | 0.6                               | 1.0          | 0.7          | 2.5          | 0.0          |
| No response       |              |                                   |              |              |              |              |
| Blank*            | 8.1          | 10.0                              | 8.5          | 9.4          | <b>13.8†</b> | 1.8          |
| No strategy*      | <b>33.8§</b> | <b>30.3§</b>                      | <b>22.3‡</b> | <b>31.6§</b> | <b>26.2‡</b> | <b>25.2‡</b> |

Note—Bold type indicates a significant prediction. NART, National Adult Reading Test; ROT, performance on the mental rotation test; COG, cognitive estimates test; FAS, FAS verbal fluency test; Wiscp, perseverative errors on the Wisconsin Card Sort Test. \*Data were log-transformed to remove skewing prior to analysis. †*p* < .05; ‡*p* < .01; §*p* < .001.

**Table 5**  
**Variance in Likelihood of R Responses Given That Elaborative Encoding Had Been Used, and Likelihood of Same Details Given at Encoding and Recognition, Explained by Age, and by Age After Psychometric Test Performance Has Been Controlled**

| Measure of Recollection | Age Alone               | Age After Controlling For |                         |                         |     |        |
|-------------------------|-------------------------|---------------------------|-------------------------|-------------------------|-----|--------|
|                         |                         | NART                      | ROT*                    | COG                     | FAS | Wiscp* |
| Words P(R)              | 3.4                     | 4.6                       | 2.0                     | 4.4                     | 1.8 | 2.5    |
| Nonwords P(R)           | 5.5                     | 8.3                       | 3.8                     | 8.3                     | 2.8 | 2.2    |
| All P(same)             | <b>21.5<sup>‡</sup></b> | <b>16.7<sup>‡</sup></b>   | <b>11.8<sup>†</sup></b> | <b>16.7<sup>‡</sup></b> | 6.8 | 9.2    |

Note—Bold type indicates a significant prediction. NART, National Adult Reading Test; ROT, performance on the mental rotation test; COG, cognitive estimates test; FAS, FAS verbal fluency test; Wiscp, perseverative errors on the Wisconsin Card Sort Test. \*Data were log-transformed to remove skewing prior to analysis. <sup>†</sup> $p < .05$ ; <sup>‡</sup> $p < .01$ .

courages focus on particular aspects of the items that would be absent if a single list were used. It is known that items that stand out lead to better memory (Von Restorff, 1933), and this may also have contributed to the pattern.

We attempted to understand why older adults show less recollective experience by conducting a series of regression analyses. The first analysis demonstrated that reported recollective experience is related to age and that the frontal measures were not successful in removing the age predictions of recollective experience. The only exception was that perseverative errors on the Wisconsin card sort and FAS fluency both removed the age effect on hit rates for words. Thus on only 2 out of 18 occasions did a frontal test remove the age effect. On the other hand, the mental rotation test removed the age effect on all measures of recollective experience for words and yielded reduced false positives for nonwords. Thus the non-frontal measures together were successful on 4 out of 12 occasions.

The second analysis showed that older adults used less elaborative encoding for nonwords and were more likely to say that they could not think of an encoding strategy than were younger adults. Here there was little evidence that any individual difference measure was successful in removing the age effect. The only measure that did so was FAS fluency, which removed the age difference on the likelihood of using deep encoding strategies for nonwords.

The third regression analysis showed that once the differences in encoding strategy were accounted for by focusing only on those items that had received deep encoding, there were no age effects on reported rates of recollective experience at all, and no clear influence of either frontal or nonfrontal test performance. This pattern is consistent with that of Perfect et al.'s (1995) Experiment 2, in which no age differences in recollective experience were found after depth-of-processing instructions were given to constrain the nature of encoding. The conclusion we draw from these three analyses is that older adults show reduced recollective experience because of poorer encoding rather than because of a failure to reintegrate item and context at retrieval. Further, we conclude that the *reduction* of age effects in recollective experience is not related to frontal test performance.

The data thus appear to favor an encoding hypothesis, although not the version outlined in the introduction. Initially, we hypothesized that older adults would show less elaborative and more maintenance rehearsal. In line with the frontal hypothesis, we expected that age differences in encoding would be removed by controlling for measures of frontal lobe integrity. We found neither of these things. We found less elaborative rehearsal (for nonwords), no age differences on maintenance rehearsal, but an age increase in no rehearsal at all. This pattern suggests that older adults are aware that maintenance rehearsal is a poor strategy, and, like younger adults, avoid its use. However, the older adults appear less able to generate appropriate elaborative strategies and are more often left without any strategy at all. Frontal test performance did not appear to be related to strategy use.

There are now several lines of evidence in favor of an encoding-based account of the deficit in recollective experience in old age. Mäntylä (1993) showed that older adults used more stereotypic encoding in a paired recall task. Perfect et al. (1995) showed that under directed encoding conditions, the age difference in recollective experience was reduced. The data here show that (1) there are age differences in encoding, and (2) once encoding differences are removed (by focusing exclusively on those items that received elaborative encoding), there is no age deficit in recollective experience.

It might be argued that we did not offer an adequate test of the frontal hypothesis, and that more, or alternative, tests of frontal functioning would have provided evidence. This is a possibility. However, our selection of frontal tests was based on the original work conducted to generate the frontal hypothesis, namely the Parkin and Walter (1992) study. Further, we added a third measure—cognitive estimation—that has been shown to be impaired in frontal patients. Clearly, adding more and more alternative “frontal” tests is not desirable, since this inflates the likelihood of finding spurious associations and increases the demands on volunteers. Our task selection was therefore principled, and the appropriate conclusion to draw is that our data do not support the frontal hypothesis as currently formulated. An alternative criticism might be that we should have combined the frontal measures to form a single “frontal” factor. There are two reasons why we did not do this. First, we wanted to be able to compare our findings with those of Parkin and Walter, who did not combine their measures. Second, and more importantly, such a step would oversimplify the nature of the role of frontal lobes in memory. There are several views on the role of the frontal lobes, but there is agreement that the frontal cortex is not a monolithic frontal system, but consists of many subsystems, each with a different role (Boller & Grafman, 1994). In this light, it is more sensible to examine *which* frontal tests effectively measure age-related decline in recollective experience, rather than to combine them as if the frontal lobes performed only one function. However, since there was no evidence for any frontal role in age-related decline, this line of reasoning was not pursued.

It is important to consider the nature of the regression analyses that have been conducted: We have attempted to see whether individual differences in frontal test performance can explain age-related changes in recollective experience. The fact that the frontal tests do not remove the age-related variance in recollective experience does not mean that frontal test performance is unrelated to recollective experience. Quite the contrary. If one correlates the frontal measures directly with measures of recollective experience, one gets correlations up to .50. Thus our data are not out of line with those of Parkin and Walter (1992). However, these correlations do not relate to the age-related *change* in performance, and hence cannot be used as a causal explanation of the age-related reduction in recollective experience. Unfortunately, Parkin and Walter did not analyze their data in a manner that would allow this hypothesis to be tested.

Initially we were concerned that asking the subjects to use a think-aloud protocol would alter the nature of the encoding they used. Since we cannot know how the group would have encoded if we had not asked them, this is an untestable hypothesis. Nonetheless, when we ran a second group of older adults using the same materials, we found performance levels on the recognition test that were indistinguishable from those of the older group, which had used the think-aloud protocol. Thus, although the paradigm is different from that which was previously used, we are reasonably confident that the findings are sound enough to support conclusions about data gained from the traditional R/K paradigm.

Clearly the coding scheme used for the encoding strategies was not very fine grained. Initially we intended to examine strategy use in greater detail and developed a coding scheme with many more categories. However, this coding scheme did not allow us to conduct the desired regression analyses; with so many categories, the frequency of use of many categories was at floor level. Thus, we favored the broader sweep because it would give us the general pattern of encoding strategy use. Future research may enable more fine grained analysis to determine more precisely the encoding changes that occur with increased age. Nonetheless, within these broad categorizations, the data were reasonably clear: Older adults were more likely to fail to come up with a strategy at all, and less likely to use elaborative encoding with nonwords. These two facts may be related: There is a semantic component to all words that is available to use for encoding with minimal effort (e.g., when one is presented with the word *table* under instructions to remember, an image of a table or an associated concept comes readily to mind). However, for nonwords there is no given semantic component, and to elaborate on a nonword requires the extra conceptual step of moving from the nonword to a related word. This may not be easy for some nonwords. For example, it is not intuitively obvious that an image or associated idea will readily spring to mind for the nonword *frumstle*. Given that older adults often report that they cannot think of a strategy to use, it is perhaps not surprising that they are also poorer in the

nonword condition, in which the appropriate strategy is least obvious.

In the introduction, we discussed the different models that have been used to interpret R and K responses. Most of the debate between advocates of different models has focused on the interpretation of K responses. Here we have focused primarily on recollective experience, and we have measured this both in terms of raw scores and in terms of the *A'* measure recommended by Donaldson (1996). Our intention was to see whether the two models would lead to different interpretations of the age-related reduction in recollective experience. They did not. Like the raw data, the measures of *A'* were related to age, and the age effects remained after we controlled for measures of frontal or non-frontal test performance. Thus we are reasonably confident that our analyses hold up regardless of the assumptions underlying the interpretation of R responses.

In summary, we sought to explain why older adults report less recollective experience. We decided to take a direct approach to the issue of how older adults encode by recording what it was they thought of when asked to encode. Our evidence suggests that reduced recollective experience in old age is not a retrieval problem, but is due to encoding deficits, particularly with items that do not offer an obvious encoding strategy. We acknowledge that the work here is preliminary, and that other categorizations of the encoding strategies might have been possible. However, we feel that our broad categories offer the most robust measures of encoding quality and have provided an interesting insight into the experiential changes in memory in old age.

## REFERENCES

- BENTON, A. L., HAMSHER, K., VARNEY, N., & SPREEN, O. (1983). *Contributions to neuropsychological assessment*. New York: Oxford University Press.
- BOLLER, F., & GRAFMAN, J. (1994). *The handbook of neuropsychology* (Vol. 9). Amsterdam: Elsevier.
- CRAIK, F. I. M., & JENNINGS, J. M. (1992). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 51-110). Hillsdale, NJ: Erlbaum.
- DONALDSON, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition*, **24**, 523-533.
- FRANCIS, W. N., & KUČERA, H. (1982). *Frequency analysis of English usage*. Boston: Houghton Mifflin.
- GARDINER, J. M., GAWLIK, B., & RICHARDSON-KLAVEHN, A. (1994). Maintenance rehearsal affects knowing, not remembering; elaborative rehearsal affects remembering, not knowing. *Psychonomic Bulletin & Review*, **1**, 107-110.
- GARDINER, J. M., & JAVA, R. I. (1990). Recollective experience in word and nonword recognition. *Memory & Cognition*, **18**, 23-30.
- GARDINER, J. M., & PARKIN, A. J. (1990). Attention and recollective experience in recognition memory. *Memory & Cognition*, **18**, 579-583.
- HORN, J. L., & CATELL, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, **26**, 107-129.
- JONES, G. V. (1987). Independence and exclusivity among psychological processes: Implications for the structure of recall. *Psychological Review*, **94**, 229-235.
- KORIAT, A., & GOLDSMITH, M. (1996). Memory metaphors and the real-life/laboratory controversy: Correspondence versus storehouse conceptions of memory. *Behavioral & Brain Sciences*, **19**, 167-228.
- LIGHT, L. L. (1991). Memory and aging: Four hypotheses in search of data. *Annual Review of Psychology*, **42**, 333-376.
- LOFTUS, G. R., & KALMAN, H. J. (1979). Encoding and use of detail in-

- formation in picture recognition. *Journal of Experimental Psychology: Human Learning & Memory*, **5**, 197-211.
- MÄNTYLÄ, T. (1993). Knowing but not remembering: Adult age differences in recollective experience. *Memory & Cognition*, **21**, 379-388.
- NELSON, H. E. (1976). A modified card sorting test sensitive to frontal lobe defects. *Cortex*, **12**, 313-324.
- NELSON, H. E. (1982). *The national adult reading test (NART): Test manual*. Windsor, U.K.: NFER-Nelson.
- PARKIN, A. J., & LAWRENCE, A. (1994). A dissociation in the relation between memory tasks and frontal lobe tests in the normal elderly. *Neuropsychologia*, **32**, 1523-1532.
- PARKIN, A. J., & WALTER, B. M. (1992). Recollective experience, normal aging, and frontal dysfunction. *Psychology & Aging*, **7**, 290-298.
- PERFECT, T. J., MAYES, A. R., DOWNES, J. J., & VAN EIJK, R. (1996). Does context discriminate recollection from familiarity in recognition memory? *Quarterly Journal of Experimental Psychology*, **49A**, 797-813.
- PERFECT, T. J., WILLIAMS, R. B., & ANDERTON-BROWN, C. (1995). Age differences in reported recollective experience are due to encoding effects, not response bias. *Memory*, **3**, 169-186.
- RAJARAM, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition*, **21**, 89-102.
- SALTHOUSE, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Erlbaum.
- SHALLICE, T., & BURGESS, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, **114**, 727-741.
- SHALLICE, T., & EVANS, M. E. (1978). The involvement of the frontal lobes in cognitive estimation. *Cortex*, **14**, 294-303.
- SHIMAMURA, A. P. (1995). Memory and frontal lobe function. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 803-813). Cambridge, MA: MIT Press.
- TULVING, E. (1985). Memory and consciousness. *Canadian Psychology*, **26**, 1-12.
- VON RESTORFF, H. (1933). Über die Wirkung von Bereichsbildungen im Spurenfeld [On the effects of formations in the trace field]. *Psychologische Forschung*, **18**, 299-342.
- YONELINAS, A. P., & JACOBY, L. L. (1995). The relation between remembering and knowing as bases for recognition: Effects of size congruency. *Journal of Memory & Language*, **34**, 622-643.

#### NOTES

1. The notion that older adults are impaired at encoding is not new (see Craik & Jennings, 1992, for a review). However, this putative encoding deficit has not before been directly linked to recollective experience.
2. Examination of the data prior to the regression analyses showed that the distributions of several variables were skewed. These variables were transformed using natural logarithms, and this significantly reduced skewing in each case. The variables concerned are marked with an \* symbol in each table.
3. An encoding strategy was classified as elaborative if any reference to semantic knowledge or episodic memory was made, including interitem associations. The classification of shallow encoding was used if item repetition or simple rhyming was used. If a subject used both kinds of encoding for an item, it was classified as elaborative.

(Manuscript received July 15, 1996;  
revision accepted for publication November 15, 1996.)