Aging effects on recognition memory when encoding and strategy are controlled

DAVID DORFMAN, MURRAY GLANZER, and JOSEPH KAUFMAN
New York University, New York, New York

It has been asserted that differences between young and old subjects in memory tasks can be eliminated if three conditions are met: (1) adequate retrieval information is given; (2) encoding is equalized for both groups; and (3) strategy differences are controlled. Although all three conditions are met in the present experiment, the age-related memory deficit still remains. It is concluded that the deficit cannot be eliminated by any of the conditions noted above.

The present study is concerned with the age-related decline in human memory and the extent to which it can be modified by study and test conditions. Craik (1977) has argued that if retrieval requirements are held to a minimum and encoding is equalized for young and old subjects, the differences between performances of young and old subjects on memory tasks should be minimal. In the experiment presented here, retrieval requirements are minimized by using recognition as the measure of memory. Encoding is equalized by using an incidental learning task to expose the stimuli to the subjects. The specific task used, lexical decision, serves to induce the same type of processing in the young and old subjects. Moreover, the task is one that the old subjects perform at least as well as the young subjects, as will be seen below.

The experimental arrangements described above were also used by Clark (1981) in an unpublished study and led to results contrary to Craik's (1977) thesis. Clark's results, however, could be questioned on the grounds that encoding was not actually equalized because the two subject groups may have had different response strategies in performing the lexical decision task. Specifically, old subjects might perform more cautiously than young subjects (Botwinick, 1978), and this may in turn affect their memory performance. In other words, the young and old subjects may place different emphasis on accuracy versus speed during the encoding task. In order to evaluate the role of this factor on aging effects in memory, the emphasis on speed versus accuracy was put under experimental control by stressing speed or accuracy for different groups of subjects.

The combinations of experimental conditions used here will provide answers to several questions concerning memory deficits that appear with aging. These questions concern whether the deficits can be removed by any one or any combination of the following conditions: (1) the minimization of retrieval (i.e., by the use of recognition);

(2) the equalization of encoding (i.e., by an incidental learning task); and (3) the specification of the speed versus the accuracy level during encoding.

METHOD

Subjects

Thirty young subjects (mean age = 19.6 years) and 24 old subjects (mean age = 74.0 years) participated in the experiment. The young subjects were New York University (NYU) undergraduates who served to fulfill a course requirement. The old subjects were volunteers from local community agencies. Only old subjects with normal memory as indicated by their scores on the Guild Memory test (Gilbert, Levee, & Catalano, 1974) were included in the study.

Apparatus and Materials

The stimuli were presented on a video display controlled by an Apple II+ computer, which also recorded subjects' responses and reaction times. Forty-eight words and 48 nonwords were presented in a lexical decision task. Of the words, 24 were high frequency (150 or more per million) and 24 were low frequency (2-8 per million), according to the Kučera and Francis (1967) word count. Within each frequency range, 12 words were one syllable long and 12 were two syllables long. The nonwords were created by changing one or two letters in 48 words matched for frequency and syllable length with the words referred to above. Data collected from a previous study with NYU undergraduates were used to assure that none of the words or nonwords were likely to give rise to incorrect lexical decisions.

The 48 distractors for the recognition task were matched for syllable length and frequency with the words used in the lexical decision task. Practice items (24 for lexical decision, 2 for recognition) were selected in the manner described above.

Design

For the lexical decision task, a $2 \times 3 \times 2$ split-plot factorial was used (age × stimulus class × speed vs. accuracy instructions). For the recognition test, a $2 \times 2 \times 2$ split plot was used (age × word frequency × speed vs. accuracy instructions). The second factor in both designs was within-subjects; the other two factors were between-subjects.

Procedure

Subjects were run individually and performed at their own pace both the lexical decision task and surprise recognition test. They made all responses by pressing buttons.

Subjects in the speed condition were told to respond as fast as possible and not to worry about mistakes. Subjects in the accuracy condition were told to respond as accurately as possible and not to worry about the amount of time they took. After each trial, subjects in the speed condition received their response time as feedback, and subjects in the

---

The authors' mailing address is: Department of Psychology, New York University, 6 Washington Place, New York, NY 10003.

Table 1
Mean Lexical Decision Response Times (in Milliseconds) and Percent Correct

	Speed Instructions		Accuracy Instructions	
	RT	% Correct	RT	% Correct
Young				
High-Frequency Words	522	94.5	815	99.7
Low-Frequency Words	600	80.5	891	94.7
Nonwords	646	86.1	1,285	96.4
Old				
High-Frequency Words	651	97.2	848	99.0
Low-Frequency Words	720	94.1	1,038	97.9
Nonwords	863	96.5	1,230	98.1

accuracy condition received feedback as to whether or not their response had been correct. All feedback was presented on the video display for 1,000 msec.

A surprise recognition test immediately followed the lexical decision task. It took 2–3 min to give the instructions for the recognition test. The list consisted of 96 words (the 48 words from the lexical decision task and the 48 distractors) and was shown one word at a time. Subjects were told that, in contrast to the lexical decision phase of the study, speed and accuracy were equally important. No feedback was given.

In the third part of the experiment, manual response times were measured. These data do not add to the interpretation of the results reported below and will not be discussed further.

During the lexical decision part of the study, the response buttons marked "yes" (for items that were words) and "no" (for items that were not words). During the recognition part, the buttons were marked "old" and "new." The "yes" ("old") button was assigned to each subject's dominant hand.

On both the lexical decision task and recognition test, each subject received a uniquely randomized list. The lexical decision task was preceded by 8 announced and 16 unannounced practice trials. The recognition test was preceded by 4 announced practice trials. The two "old" words for the recognition practice were drawn from the announced lexical decision practice list. During the announced practice phases, the experimenter remained in the room to assure that subjects knew how to perform the tasks properly. At all other times, the experimenter was outside of the room.

RESULTS AND DISCUSSION

Table 1 shows the lexical decision data: response times and percent correct for each stimulus category for all conditions. Response times for stimulus categories show the usual pattern: both young and old subjects responded fastest to high-frequency words, next fastest to low-frequency words, and slowest to nonwords. The differences in response times are significant [$F(2,100) = 67.5$, $p < .001$]. The data also show that the speed versus ac-

curacy instructions are effective for both young and old subjects. In the accuracy condition, both groups showed an increase in mean response time [$F(1,50) = 19.6$, $p < .001$] and percent correct [$F(1,50) = 52.5$, $p < .001$]. The data show that the old subjects performed the lexical decision task at least as well as the young subjects. The difference in response times between the two groups is not significant [$F(1,50) = 1.48$, n.s.], but the old subjects were significantly more accurate [$F(1,50) = 35.4$, $p < .001$]. The difference in accuracy may have resulted from the higher educational level of the older subjects. Whereas 75% of the older subjects had at least one postsecondary degree, 60% of the young subjects had completed less than 2 years of college and none of them had a postsecondary degree.

The hits and false alarm rates for the recognition test are shown in Table 2. Also shown in Table 2 are A' measures of recognition memory. A' is a nonparametric signal detection theory measure of memory sensitivity independent of response bias (Pollack, Norman, & Galanter, 1964). It can be interpreted in the same fashion as a two-alternative forced-choice test: .5 is chance performance; 1 is perfect performance. Its advantage in the present case is that it requires minimal assumptions about the underlying distributions.

Young subjects showed higher recognition performance as measured by A' across all conditions. Although the young subjects appear to be helped by accuracy instructions and the old subjects did not, the lack of a significant main effect of instruction or an age × instruction interaction do not support further consideration now. As expected, the effect of word frequency is significant [$F(1,50) = 63.5$, $p < .001$].

The main point established by this study is that despite the use of a memory task which minimizes the role of retrieval, and despite the equalizing of encoding by both the use of an incidental learning task and the specification of speed versus accuracy sets, the age-related deficit is clear. The appearance of the deficit is especially striking since the old subjects were more competent on the encoding task than were the young subjects.

REFERENCES

BOTWINICK, J. (1978). *Aging and behavior* (2nd ed.). New York: Springer.
CLARK, E. O. (1981). *Semantic and episodic memory impairment in*

Table 2
Mean A', Mean Percent Hits, and Mean Percent False Alarms on the Surprise Recognition Test

	Speed Instructions			Accuracy Instructions		
	A'	% Hits	% False Alarms	A'	% Hits	% False Alarms
Young						
High-Frequency Words	.799	74.9	33.1	.803	64.8	21.4
Low-Frequency Words	.872	81.3	20.8	.914	78.4	8.9
Old						
High-Frequency Words	.764	63.8	26.7	.736	42.8	17.4
Low-Frequency Words	.844	70.0	17.0	.840	57.0	10.4

old age. Unpublished doctoral dissertation, New York University, New York.

CRAIK, F. I. M. (1977). Age differences in human memory. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 320-384). New York: Van Nos Reinhold.

GILBERT, J. G., LEVEE, P. F., & CATALANO, F. L. (1974). *Guild memory test manual.* Newark, NJ: UNICO National Mental Health Foundation.

KUČERA, F., & FRANCIS, W. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

POLLACK, I., NORMAN, D. A., & GALANTER, E. (1964). An efficient non-parametric analysis of recognition memory. *Psychonomic Science,* 1, 327-328.