Task complexity and age differences in working memory

MARY L. GICK
Carleton University, Ottawa, Canada

and

FERGUS I. M. CRAIK and ROBIN G. MORRIS
University of Toronto, Toronto, Canada

This study investigated age-related differences in working memory using a modified version of the Daneman and Carpenter (1980) working memory task. The subjects were required to verify a series of sentences, and then at the end of each series recall the final word of each sentence. Each series varied in length from one to five sentences. Performance on this task was compared with performance in a word-alone condition, in which the subject had to remember an equivalent list of single words but without sentence verification. When sentences of positive grammatical form were used in the sentence-span condition, age differences were no greater than in the word-alone condition; however, the age decrement increased when sentences of negative grammatical form were used. There were no interactions between age and pacing or between age and the number of sentences in each set. These results are discussed in relation to theories of age differences in working memory.

In his classic monograph on aging and skilled performance, Welford (1958) provided good evidence that older people perform poorly on a wide range of perceptual-motor tasks. Welford proposed that an age-related decrement in short-term memory was the crucial factor underlying the poorer performance; in particular, he claimed that short-term retention in the elderly is more susceptible to interference from the performance of some other activity. This was illustrated, for example, in the performance of subjects of different ages on a continuous choice reaction time task in which subjects pressed the response key corresponding to the stimulus light that was on; each correct response immediately caused another light in the stimulus array to be illuminated. In the straightforward version of this task, older people showed little decrement, but they showed a much greater loss relative to young subjects when the task was complicated by making the person press the key corresponding to the previously illuminated stimulus, that is, when working one back in the series. In this more complicated version, the subjects had to retain the position of the current stimulus in short-term memory while making their response to the previous stimulus.

The necessity to manipulate information held in short-term memory while carrying out further operations on the stored items is the hallmark of "working memory" as described by Baddeley and Hitch (1974). Substantial age decrements have been found in tasks that require this type of cognitive processing. For example, the elderly typically perform poorly on a dichotic listening task that requires the person to simultaneously listen to two lists and report each back separately (Craik, 1977; Inglis & Caird, 1963). In a more recent study, Wright (1981) showed that older adults also perform poorly on the Baddeley and Hitch (1974) working memory paradigm, in which subjects must retain a short list of digits in short-term memory while simultaneously carrying out a verbal reasoning task. Age differences in this paradigm were explored further in a study parallel to the present experiment (Morris, Gick, & Craik, 1988).

Age-related deficits in memory functioning have also been explored in the context of memory for discourse. It has been suggested that older people have difficulty in understanding discourse because of a reduction in working memory capacity (Cohen, 1981; Light & Anderson, 1985; Light, Zelinski, & Moore, 1982; Spilich, 1983; Zelinski, Light, & Gilewski, 1984). The relationship between working memory functioning and text processing in the aged has been investigated by Light and Anderson (1985) using the Daneman and Carpenter (1980) sentence-span task. In this task subjects are required to read a series of sentences. After the entire set of sentences has been read, the subjects are required to report the final word of each sentence in the original order. The task constrains the subjects to simultaneously process each sentence and hold the set of final words in mind. Light and Anderson (1985) reported that older adults are impaired on this

sentence-span task, although the extent of the impairment is the same as in an ordinary word-span task.

In the present study, we modified the Daneman and Carpenter (1980) task so that subjects were obliged to process the stimulus material actively. Instead of simply reading, the subject had to decide whether the factual statement presented in each sentence was true or false, and had to respond manually using two-choice response keys; after the series of sentences was presented, the subject attempted to recall the series of final words.[1] By requiring the participant to make a decision, the task is more analogous to the Baddeley and Hitch (1974) working memory task and was expected to be more sensitive to age-related deficits in working memory. This task was used to explore two factors, complexity and pacing, that may contribute to the age-related decrement.

Working memory tasks require simultaneous processing of incoming materials and storage of the products of that processing, and sometimes require further operations on the stored products; all of these aspects can vary in complexity. There is good evidence that older people are differentially affected by increases in complexity (Cerella, Poon, & Williams, 1980; Salthouse, 1982; Welford, 1958), and this interaction between age and complexity would be expected to appear in working memory tasks. Some evidence to this effect has been presented (Light & Anderson, 1985; Spilich, 1983; Wright, 1981). In the present study, and in the parallel experiment reported by Morris et al. (1988), complexity was manipulated by varying the difficulty of the sentences to be verified; it was expected that older people would be more penalized by complex sentences and by requirements of heavier memory loads. In Light and Anderson's (1985) study, age differences were no greater on reading span than on simple word span. However, it may be necessary to engage the subjects in making a decision about the sentences in order to draw on sufficient processing resources to produce the difference.

A related aspect of the factor of complexity in working memory tasks is that of divided attention. There is some evidence to suggest that the division of attention between ongoing processing and storage functions is particularly disruptive to performance in older people (Craik, 1977; Welford, 1958). Salthouse (1982) suggested that divided attention may simply act to increase the complexity of the overall task and that the observed age differences in divided attention tasks may reflect the difficulties experienced by older people in any situation involving complex operations. One purpose of the present experiment was to further examine the question of age differences in divided attention situations. Divided attention was studied in the present experiment by comparing recall performance on the reading-span task, which requires the dual tasks of reading and remembering, to recall performance on a single task that requires only memory for words. In summary, two major reasons for carrying out the experiment were to examine possible interactions between age and task complexity in a working memory set-

ting, and to see whether division of attention acts similarly to other ways of increasing complexity in this situation.

Another factor explored in the present study was the time allowed for subjects to perform the task. There is strong evidence that older people are especially penalized when the task is paced, or when responses must be made within a given time limit (Welford, 1958, 1977). Pilot work with the present task suggested that older people performed quite well when given unlimited time to verify each successive sentence, but were substantially poorer than younger participants when a time limit was imposed.

In overview, the present experiment examined whether there are age differences in working memory, and whether such age differences are affected by the complexity of the constituent tasks and by pacing of the subjects' responses. In each trial, 1, 2, 4, or 5 sentences were presented successively. The subject's task was to judge whether each statement was true or false, and then to recall the final words from all the sentences; that is, the subject recalled a maximum of 1, 2, 4, or 5 words, in the original order, following presentation and verification of all the sentences. Task complexity was manipulated in three ways. First, sentence complexity was varied by presenting either positive sentences (e.g., "Cats usually like to hunt mice" or "A canary may often be bigger than a horse") or negative sentences (e.g., "Bookcases are not usually found by the sea" or "Children never like to play at the beach") (Chase & Clark, 1972). Complexity was also manipulated by varying the necessity to divide attention: most trials required the subject to divide attention between holding words in mind and processing the next sentence, but some trials were given in which only the set of words to be recalled in order was given. Finally, task complexity was manipulated by varying the number of sentences presented on each trial, thereby varying the memory load. In addition to these manipulations of task complexity, pacing was varied either by giving subjects unlimited time to study each sentence or by imposing a time limit of 8 sec. It was expected that older people would perform less well on this working memory task, and that the age-related decrement would be amplified by greater degrees of complexity and by the imposition of pacing.

METHOD

Subjects

Eighteen young and 18 elderly subjects were tested in the experiment. The young subjects were college students who signed up for the experiment and received $10 for their participation. The average age of the young subjects was 21.9 years (SD=2.0), and their average score on the Mill Hill Vocabulary Test was 14.6. The elderly subjects were obtained from the pool of volunteers who regularly participate in experiments and who are reimbursed for their transportation expenses. These older subjects were healthy people who live independently in the Mississauga area. The average age of these subjects was 68.1 years (SD=4.8), and their average Mill Hill score was 16.2. The vocabulary scores of the elderly subjects were reliably higher than those of the college students [$t(34) = 2.42, p < .05$]. The young subjects had an average of

14.8 years (SD=1.0) of formal education, whereas the older group had received an average of 12.5 years (SD=3.2). This age difference was also reliable [$t(34) = 2.77$]. Thus the younger group had received more formal education, but the older group scored reliably higher on the synonym vocabulary test; this pattern is typical in our experience.

Design

The design was a 2 (old/young) × 2 (paced/unpaced) × 3 (simple/complex/word alone) × 4 (1/2/4/5 words per set) experiment, with age a between-subjects factor and all other factors varied within subjects. The first within-subjects factor, pacing, refers to whether the subject was given a time limit in which to verify each sentence (paced) or was given no such time limit (unpaced). The second factor, complexity, refers to whether a sentence was presented for verification and, if so, whether it was simple (positive) or complex (negative). The third within-subjects factor set size, refers to the number of sentences presented on each trial.

In addition, a series of sentence-alone trials was given in which the subjects were required to verify the sentences but were not required to retain the final words. This series acted as a control condition for the effects of memory load on verification latency.

Materials

The experiment was designed so that each memory word was used equally often in each of the main conditions: unpaced-positive sentences, unpaced-negative sentences, paced-positive sentences, paced-negative sentences, unpaced-word alone, paced-word alone. The words were taken from a pool of 448 one- and two-syllable concrete nouns. A positive sentence and a negative sentence were constructed for each word. Half of them were true, and half were false (e.g., "A river is usually longer than a stream" and "An ocean is not larger than a stream").

The entire pool of sentences was divided randomly into 12 sets, 6 with positive sentences and 6 with negative sentences. Each set had an equal number of true and false sentences, and none of the final words was repeated within a set. Each subject was presented with a total of 6 sets, according to the six main conditions listed above, with the constraint that none of the final words was repeated for a given subject. Using the sentences presented in the previous paragraph, if a subject received the positive *river-stream* sentence, the same subject did not receive the negative *ocean-stream* sentence in any of the other conditions. Across subjects, each of the 6 sets of materials was used equally often in each of the six main conditions. Within a condition, items were randomly chosen without replacement from the appropriate set.

Procedure

The subjects were tested individually in two sessions on separate days, one for the unpaced conditions and the other for the paced conditions. The order of the sessions was counterbalanced across subjects. On each trial, sentences or words were presented sequentially on the monitor of a Commodore PET 8200 microcomputer. In the sentence-span trials, the subject read the item silently and pressed one of two keys for the verification response. In the word-alone trials, the subject was required to press one of the two response keys to signal readiness for the next word. In the word-alone and sentence-span trials, the subject's response was followed immediately by the next item. After responding to the last item, the subject was required to report the words in their original serial order. The procedure was the same for the sentence-alone trials, with the exception that for each trial there was only one sentence, with no memory requirements.

In the unpaced conditions, there was no time restriction on responding to each sentence. However, in the paced conditions, the subject had a maximum of 8 sec to verify each sentence. The elapsing time interval was indicated by an array of plus signs that incremented by one item every second. If the subject exceeded this time limit, he/she proceeded to the next item and the computer recorded an error for that trial, discarding the trial latency. Similarly, if the subject made an incorrect response, the trial latency was discarded from the analysis.

For each type of memory test (word alone, positive sentences, negative sentences) combined with each set size (1, 2, 4, and 5 items), there were four trials. For each set size, all subjects received the word-alone task first, followed by the sentence-span tasks, with positive or negative sentences. In order to control for practice, the four trials for each condition were split into two sets, with two trials of a given set size blocked together and performed in ascending order (1, 2, 4, and 5) and the other two trials in descending order (5, 4, 2, and 1). The subjects were given a rest between ascending and descending orders of the trials. The subjects were always informed of the type of trial (sentence span, word span, or sentence alone) prior to stimulus presentation. Before each set of test trials, the subjects were given a practice trial with sentences randomly chosen to be positive or negative.

In order to control for practice on the sentence-alone trials, they were mixed in with the other trials as follows: In each of the paced and unpaced sessions, four sentences were presented alone for verification. Two sentences were positive and two were negative. Two were presented prior to set sizes 1 and 4 in the ascending order of the span task, and the other two prior to set sizes 5 and 2 in the descending order. The order of the simple and complex sentences was randomly determined.

The subjects were encouraged to be as accurate as possible on the sentence verification task and were instructed that accuracy was more important than speed.

RESULTS

Sentence Verification Errors

The mean percentages of errors made on the sentence verification task are shown in Table 1. The data were analyzed by means of a four-way analysis of variance (ANOVA) involving age (young or old), pacing (paced or unpaced), sentence complexity (positive or negative), and set size (1, 2, 4, or 5 sentences). The ANOVA revealed main effects of age [$F(1,34) = 21.8, p < .001$], pacing [$F(1,34) = 27.7, p < .001$], sentence complexity [$F(1,34) = 76.7, p < .001$], and set size [$F(3,102) = 3.35, p < .05$]. That is, more errors were made by older subjects, under paced conditions, and with negative sentences. Somewhat fewer errors were made in the set size 1 condition than in the other three set size conditions. Two two-way interactions were significant: between age and sentence complexity [$F(1,34) = 14.3, p < .001$] and between pacing and complexity [$F(1,34) = 11.0, p < .005$]. That is, both age and pacing differentially increased the number of errors made on complex, as opposed to simple, sentences. No other main effects or interactions were significant; in particular, it is notable that the effect of pacing was no greater on the old than on the young subjects.

Table 1 also shows the percentages of verification errors on the sentence-alone trials, that is, trials on which the subjects were not required to recall the last words of the sentences. Errors were surprisingly high on these trials. Performance in the sentence-alone condition was compared with that in the sentence-span condition of set

Table 1
Mean Error Rates (%) and Standard Deviations for Sentence Verification on the Sentence-Span and Sentence-Alone Conditions

	Young Adults				Old Adults			
	Positive Sentences		Negative Sentences		Positive Sentences		Negative Sentences	
	M	SD	M	SD	M	SD	M	SD
Paced Conditions								
Sentence Span								
One Sentence	4.2	0.9	15.3	1.7	2.8	0.8	19.4	2.3
Two Sentences	4.9	0.7	14.6	1.1	9.0	0.9	31.3	1.7
Four Sentences	7.6	0.8	16.3	1.0	8.0	1.1	25.7	1.3
Five Sentences	6.4	0.6	13.3	1.1	9.0	0.7	30.0	1.1
Sentence Alone	19.4	2.9	16.7	2.9	11.1	2.1	27.8	3.4
Unpaced Conditions								
Sentence Span								
One Sentence	2.8	0.8	2.8	1.1	4.2	0.9	13.9	1.7
Two Sentences	2.8	0.5	6.9	1.2	4.9	0.7	19.4	1.4
Four Sentences	2.1	0.4	7.6	0.7	6.9	0.8	17.7	1.3
Five Sentences	2.8	0.4	6.1	0.5	5.0	0.5	18.3	1.5
Sentence Alone	2.8	1.2	8.3	1.9	5.6	1.6	33.3	3.3

size 1 by means of a four-way ANOVA in which age was a between-subject variable and pacing, sentence complexity, and memory (sentence alone vs. sentence span 1) were within-subjects variables. The analysis revealed the counterintuitive result that more errors were made on the sentence-alone trials, in which no memory was required [$F(1,34) = 8.57, p < .01$]. We have no convincing explanation for this result. It is possible that, due to the intermingling of these trials with sentence-span and word-alone trials, subjects had trouble switching sets. That is, they had to remember *not* to recall the words, thereby imposing an additional memory requirement that resulted in errors. The only other significant effect involving the memory factor was a pacing × complexity × memory interaction [$F(1,34) = 7.09, p < .01$], but once again we have no immediate explanation for this result.

Sentence Verification Latencies

Table 2 shows the verification latencies in the sentence-span conditions, averaged over trials in which a correct response was made by the subject. These data were analyzed using a four-way ANOVA with age (young or old) as a between-subject factor and pacing (unpaced or paced), sentence complexity (positive or negative), and set size (1, 2, 4, or 5 sentences) as within-subjects factors. This analysis yielded a main effect of age [$F(1,34) = 4.61, p < .05$], showing that the older subjects took somewhat longer to respond. The latencies were significantly decreased in the paced condition [$F(1,34) = 62.0, p < .001$], suggesting that the time constraint made all subjects respond more rapidly. Sentence complexity significantly increased verification latencies [$F(1,34) = 172.1, p < .001$], as did set size [$F(3,102) = 35.2$,

Table 2
Mean Latencies (in sec) and Standard Deviations for Sentence Verification on Sentence-Span and Sentence-Alone Conditions for Young and Old Subjects

	Young Adults				Old Adults			
	Positive Sentences		Negative Sentences		Positive Sentences		Negative Sentences	
	M	SD	M	SD	M	SD	M	SD
Paced Conditions								
Sentence Span								
One Sentence	3.33	0.60	4.76	0.73	4.20	0.79	6.02	1.49
Two Sentences	3.83	0.63	4.95	0.77	4.67	1.40	5.76	1.10
Four Sentences	4.09	0.42	5.21	0.67	5.00	1.14	6.02	0.88
Five Sentences	4.33	0.54	5.63	0.66	4.90	1.10	6.23	0.86
Sentence Alone	4.17	0.99	5.01	0.96	4.40	1.35	5.63	1.77
Unpaced Conditions								
Sentence Span								
One Sentence	4.45	1.16	6.85	2.11	5.32	1.52	7.61	2.80
Two Sentences	4.79	1.16	8.07	3.09	5.36	1.33	7.85	1.96
Four Sentences	5.93	1.64	8.35	2.17	7.09	2.40	9.83	3.79
Five Sentences	6.29	1.90	9.17	2.70	7.19	2.03	9.95	3.33
Sentence Alone	4.62	2.20	6.05	2.01	5.26	2.72	6.88	2.65

$p < .001$]. Subjects took longer to verify sentences that were complex and that were incorporated into large sets of sentences, indicating that a large memory load interfered with sentence verification. There were significant interactions between pacing and sentence complexity [$F(1,34) = 34.6, p < .001$] and between pacing and set size [$F(3,102) = 11.2, p < .001$]; in both cases, pacing reduced the effect of the other variable. There were no other significant interactions, including those between age and pacing and between age and sentence complexity.

The latencies for the sentence-alone conditions, in which retention of the final words was not required, are also shown in Table 2. The sentence verification latencies obtained with the sentence-alone trials were compared with those obtained for span trials of set size 1 in order to determine the effects of adding a memory load to verification. A four-way ANOVA involving age, pacing, sentence complexity, and memory (sentence alone vs. sentence span 1) indicated that latencies were no different between the sentence span 1 and sentence-alone tasks ($F < 1$). Furthermore, there were no significant interactions involving the memory factor. Overall, then, the additional requirement of recalling the final word after verifying a sentence did not impair verification latency for either old or young people.

The data and analyses on sentence verification errors and latencies are secondary to the main focus of the experiment, which was on recall performance. However, the analyses have shown that the older participants made more errors and took longer to verify sentences in the reading-span task than did their younger counterparts; thus a finding of an age decrement in recall cannot readily be attributed to the fact that older subjects completely traded off performance on recall for performance on sentence verification, and thereby did extremely well on the verification task. However, it is not possible to say exactly how younger and older people allocate effort to the two tasks. It is noteworthy that in this case pacing was not differentially disruptive to the verification task for older subjects.

Recall Performance

In order to obtain a valid measure of the memory aspect of the working memory task, uncomplicated by errors in sentence verification, analysis of the recall data was first based on those trials in which all sentences were verified correctly. These data are shown in Table 3. The table also shows recall in the words-alone condition (i.e., the condition in which only single words were presented for standard serial recall).

Two major analyses were carried out on the proportions of words recalled correctly, based on the raw data in Table 3. First, the effects of sentence complexity were examined by comparing performance in the two conditions involving sentence verification (positive and negative sentences); second, the effects of division of attention were examined by comparing the words-alone condition with the positive-sentence condition. In both cases, set size 1 data were not included, since there are very clear ceiling effects in this condition.

In the first analysis, a four-way ANOVA was carried out on the data from the conditions with positive or negative sentences. In this analysis, age was a between-subjects factor, and pacing, sentence complexity, and set size (2, 4, or 5 sentences) were within-subjects factors. There was a main effect of age [$F(1,34) = 13.0, p < .001$], showing that older subjects recalled fewer final words. There were also main effects of pacing [$F(1,34) = 22.1, p < .001$], complexity [$F(1,34) = 89.8, p < .001$], and set size [$F(3,68) = 119.8, p < .001$]. These effects indicate decreased recall under paced conditions, with negative sentences, and with larger set sizes, respectively.

The interaction between age and pacing was not significant ($F < 1$), showing, surprisingly, that older people were not differentially penalized by pacing in this situation. However, there were significant interactions between age and complexity [$F(1,34) = 14.9, p < .001$], between pacing and complexity [$F(1,34) = 4.78, p < .05$], and between pacing and set size [$F(2,68) = 7.48, p < .001$]. These interactions show, respectively,

Table 3
Mean Memory Scores and Standard Deviations for Young and Old Adults for Trials
with No Verification Errors

	Young Adults						Old Adults					
	Positive Sentences		Negative Sentences		Words Alone		Positive Sentences		Negative Sentences		Words Alone	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Paced Conditions												
One Sentence	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	0.92	0.16	1.00	0.00
Two Sentences	1.96	0.09	1.79	0.55	1.99	0.06	1.89	0.29	1.15	0.61	1.93	0.23
Four Sentences	2.74	0.78	1.83	1.01	3.81	0.21	2.53	0.96	1.22	0.98	3.64	0.55
Five Sentences	2.67	0.94	2.15	1.27	4.15	0.46	2.22	1.06	0.75	1.09	3.65	0.64
Unpaced Conditions												
One Sentence	1.00	0.00	0.99	0.06	1.00	0.00	1.00	0.00	0.97	0.11	1.00	0.00
Two Sentences	1.95	0.13	1.78	0.32	2.00	0.00	1.85	0.22	1.51	0.42	2.00	0.00
Four Sentences	2.99	0.78	2.82	1.01	3.86	0.17	2.58	0.76	1.81	1.14	3.63	0.42
Five Sentences	3.57	0.87	2.92	1.09	4.42	0.38	3.25	1.09	2.08	1.60	4.15	0.69

that older subjects were relatively more penalized by negative sentences, that negative sentences were associated with particularly low levels of recall under paced conditions, and that the proportion of final words recalled decreased with increasing set size, particularly under paced conditions.

A second analysis was conducted to compare the performance on the word-alone and sentence-span conditions. To investigate the effect of combining remembering with sentence verification, without the added complication of complexity, the comparison was made between word-alone trials and sentence-span trials with positive sentences. The data were analyzed using a four-way ANOVA with age (young or old) as a between-subjects factor and pacing (paced or unpaced), divided attention (word alone or sentence span), and set size (2, 4, or 5) as within-subjects factors. This analysis revealed main effects of age [$F(1,34) = 5.83, p < .05$], pacing [$F(1,34) = 12.5, p < .001$], divided attention [$F(1,34) = 90.5, p < .001$], and set size [$F(2,68) = 155.6, p < .001$], showing that recall was lower for older subjects, under paced conditions, with the sentence-span task, and with larger set sizes, respectively. The interactions between age and pacing and between age and divided attention were not significant (both Fs < 1). That is, age differences were not increased by pacing the subjects or by combining remembering with sentence verification. The only significant interactions in the analysis were between pacing and set size [$F(2,68) = 16.0, p < .001$] and between set size and divided attention [$F(2,68) = 31.5, p < .001$]. These interactions showed that the decline in recall performance with increasing set size was greater under paced conditions and with the sentence verification task, respectively.

Some of these effects can be seen more easily in Figure 1, in which mean levels of recall are shown as a function of age, experimental condition, and set size. The data are combined over paced and unpaced conditions, and are shown as proportions of the presented set size in each case. There are evident ceiling effects in the set size 2 conditions, especially for the word-alone task. However, if these points are discounted, the striking feature of the figure is the lack of interactions with set size: neither age nor experimental condition amplify the set size effect. The figure also shows that the age difference is relatively constant between the word-alone conditions and the positive-sentence conditions (i.e., no interaction between age and divided attention), but that the age difference increases substantially as a function of sentence complexity (i.e., an interaction between age and positive/negative sentences). In summary, although we predicted that age decrements in working memory performance would be exacerbated by any increase in difficulty of the component tasks, this prediction was not borne out by the results. Older people recalled fewer words, but this decrement was amplified only by increasing the complexity of the sentence to be verified, not by the simple presence of the sentence verification task, and not by increases in set size.
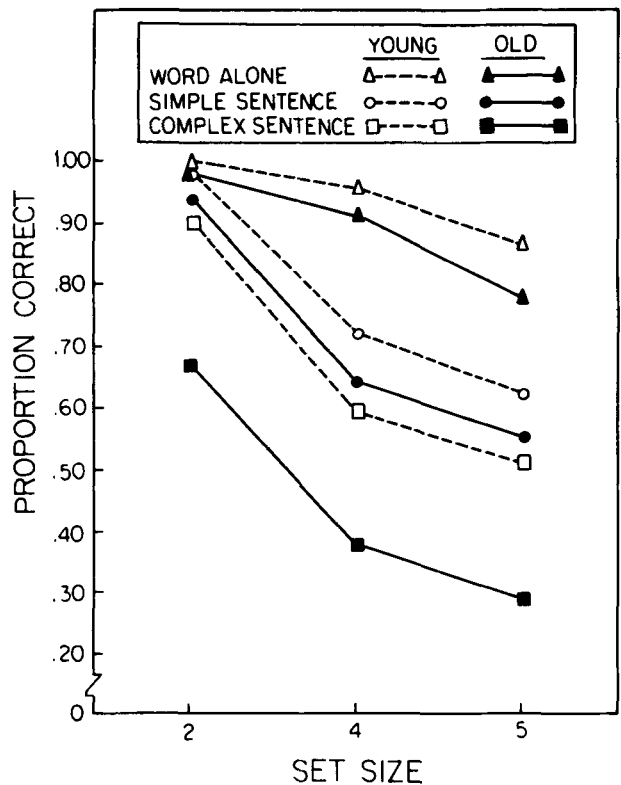

Figure 1. Proportions of words recalled as a function of age, set size, and experimental condition (excluding trials on which verification errors were made). Data are collapsed over paced/unpaced conditions.

Apparently, only certain types of difficulty give rise to special problems in the elderly—in this task at least.

It is also of interest to analyze recall data from trials on which at least one sentence verification error was made. If subjects trade off the verification and recall tasks, it might be expected that recall scores would be higher when verification errors were made. It might also be expected that such variables as sentence complexity and set size would have reduced effects on recall if attention was more on the recall task. On the other hand, Hitch and Baddeley (1976) reported no trade-offs of this type, finding instead that good recall performance in a working memory task was associated with faster response times on a concurrent verbal reasoning task. Table 4 shows both the average proportions of words recalled and the absolute number of words recalled on trials containing sentence verification errors. The data are given for set sizes 4 and 5 only, since there were few errors at smaller set sizes. In addition, since there were so few relevant observations, the data were collapsed over the pacing variable and combined over pairs of adjacent subjects to form 9 macrosubjects in each age group for the purpose of analysis.

The overall mean number of words recalled by older subjects on trials with errors was 2.57; the mean was 2.05 on trials without errors (from Table 3). Thus for older

Table 4
Words Correctly Recalled on Trials on Which Sentence
Verification Errors Were Made

| | Set Size 4 | | Set Size 5 | |
	Simple	Complex	Simple	Complex
	(a) Average Proportions of Words			
Young Adults	.65 (.07)	.66 (.13)	.55 (.20)	.57 (.11)
Old Adults	.67 (.23)	.60 (.13)	.52 (.14)	.52 (.18)
	(b) Average Number of Words			
Young Adults	2.60 (.30)	2.64 (.51)	2.75 (1.01)	2.85 (.54)
Old Adults	2.68 (.90)	2.40 (.50)	2.60 (.69)	2.60 (.88)

Note—Standard deviations are given in parentheses.

subjects there is some evidence of a trade-off between memory and verification since more words were recalled on trials with errors. However, this evidence was absent in the younger subjects, who recalled an average of exactly 2.71 words in both trials with errors (Table 4b) and trials without errors (Table 3, set sizes 4 and 5). A three-way ANOVA (age × set size × sentence complexity) on the data in Table 4a yielded only one reliable effect—a main effect of set size [$F(1,16) = 25.1, p < .001$] (i.e., the proportions of words recalled were lower from set size 5 than from set size 4). However, an ANOVA on absolute number of words recalled (Table 4b) yielded no significant effects. Apparently, then, the number of words recalled remained constant across conditions in trials with errors, and the set size effect found with proportions (Table 4a) appears to have resulted from a constant number being expressed as a proportion of 5 rather than of 4.

It thus appears that the effects of aging and of sentence complexity on recall that appear in trials without verification errors (Table 3 and Figure 1) are eliminated in trials with verification errors (Table 4). This result suggests that on some trials the greater difficulty associated with negative sentences and with increased age is "absorbed" by the verification task and shows up as an error. In these cases, there is no further effect of aging or complexity on recall scores.

DISCUSSION

In summary, the experiment showed a substantial age-related decrement in working memory performance, confirming the results of other studies (Light & Anderson, 1985; Morris et al., 1988; Spilich, 1983; Wright, 1981). However, the main purpose of the study was not simply to demonstrate an effect of aging, but to explore the effects of other variables on the age-related deficit. In particular, the effects of pacing, divided attention, and various types of task complexity were examined. The major focal point of the experiment was the effect of divided attention on performance in younger and older subjects, and the similarity (if any) of the effects of divided attention to those of other types of complexity, namely syntactical complexity and set size. Divided attention was ex-

amined in the sentence verification task by comparing the sentence-alone condition with the sentence span 1 condition (in which subjects also had to recall the final word of the sentence). This comparison showed no effects of division of attention in the verification latency data, and the counterintuitive result for error data that divided attention reduced errors (i.e., there were fewer errors in the sentence span 1 condition). In neither case did age interact with divided attention. In the recall data, division of attention was examined by comparing the word-alone and the positive-sentence conditions. In this case, the addition of the verification task did reduce recall performance, but the effect did not interact with age. This lack of interaction between age and word span/sentence span was also reported by Light and Anderson (1985). The present result also confirms the absence of an interaction between age and divided attention in working memory paradigms reported by Wright (1981), by Baddeley, Logie, Bressi, Della Sala, and Spinnler (1986), and by Morris et al. (1988).

Given the many previous studies that have reported interactions between age and divided attention (e.g., McDowd & Craik, in press; Salthouse, Rogan, & Prill, 1984; Welford, 1958), the apparent discrepancy between the two sets of results is quite puzzling. Morris et al. (1988) suggested that the age × divided attention interaction does not appear when one task involves the continuous rehearsal of span or subspan lists of verbal items, but further work is required to corroborate this suggestion. For the moment, it may be noted that division of attention is not inevitably more detrimental to the performance of older people than to that of younger people (Wright, 1981).

How do the effects of divided attention compare with the other two types of complexity used in the experiment? Both sentence complexity and set size affected performance on the present task. Verification errors increased, verification latencies increased, and recall decreased as sentence complexity and set size increased. The factors differed, however, in their interactions with age. Aging exacerbated the effects of sentence complexity (for verification errors and for recall scores), but the effects of age and set size did not interact in any measure. The experiment thus yielded the interesting result that some, but not all, variables that affect working memory performance have exaggerated effects on older people. Increases in sentence complexity did affect older subjects more than younger subjects, but neither set size nor division of attention had a differential effect on the two age groups. Obviously, not all sources of task difficulty or types of complexity are equivalent, and differential interactions with age may provide a method to analyze the nature of "complexity" in greater detail (see also Cerella et al., 1980; Salthouse, 1982).

A second focus of the experiment was the effects of pacing on working memory performance: subjects were paced by having a time constraint placed on sentence verification. This manipulation had the effect of speeding up

verification latencies at the expense of memory performance; however, contrary to prediction, pacing had no differential effect on performance of the older group. The age × pacing interaction was not statistically reliable for any of the three dependent variables—recall, verification latency, or verification errors. Given the strongly detrimental effects of pacing on older people reported by previous investigators (e.g., Welford, 1958, 1977), it can only be concluded that the present combination of intelligent, active subjects and relatively benign pacing was not sufficient for effects to be seen. It appears, then, that older people are not automatically more heavily penalized by any variable that reduces performance. However, pacing did exacerbate the deleterious effects on performance of sentence complexity and increasing set size.

The finding that age interacts with some, but not all, variables that affect working memory performance agrees with a similar pattern of results reported by Morris et al. (1988). In that study, performance of younger and older adults was examined using the Baddeley and Hitch (1974) working memory paradigm. Subjects were first given a short list of words to remember, they were then given a single sentence to verify, and finally they recalled the words. Set size was varied by giving a memory preload of either two or four words, and the complexity of the sentences to be verified was again manipulated by presenting positive or negative sentences. Recall scores were noninformative in the Morris et al. study, since subjects were asked to rehearse aloud continuously, and thus very few errors were made. In the sentence verification latency data, Morris et al. found effects of age, sentence complexity, and set size; the age × set size interaction was not reliable. These findings all parallel the results from the present study. Similarly, both studies found reliable effects of sentence complexity and of set size, but no age × set size interaction in the verification error data. The two differences in results were, first, a reliable main effect of age on verification errors in the present study but not in that of Morris et al. and, second, a significant age × sentence complexity interaction in the latency data of Morris et al. but not in the present data. However, the age × complexity interaction was reliable in the *error* data of the present study. Our claim, therefore, is that sentence complexity interacted with age in both studies, as shown in the error data of the present experiment but in the latency data of the Morris et al. study.

Taken together, the present results and those of Morris et al. (1988) do not support any hypothesis that postulates a general deficiency in information-processing capabilities in the elderly. For example, Craik and Byrd (1982) attributed age decrements in performance to a reduction in processing resources; however, it seems reasonable by this view that any variable that reduces performance in the young adult group should have a greater effect in the elderly (i.e., that a general lack of resources should exacerbate any form of difficulty). Similarly, Salthouse (1982) suggested that age-related deficits in working memory arise from the added complexity of combining two tasks. In other words, small decrements in component mental operations are compounded when they are combined. However, this theoretical viewpoint also appears to predict age-related interactions with any variable that decreases performance. Neither theory appears to give a satisfactory account of the present data.

It seems that not all forms of difficulty or complexity are equivalent, and that not all such variables present equivalent problems to older people. The researcher's task is presumably to analyze in finer detail the types of processing demanded by various manipulations, and to discover which types of complexity present special difficulties to the elderly. One suggestion that may fit both the present results and those of Morris et al. (1988) is that aging effects the efficiency of the *processing* aspects of working memory tasks, rather than the passive holding aspects; that is, aging affects the ability to verify complex sentences in working memory situations, without having much effect on the ability to rehearse items held in the articulatory loop (Baddeley & Hitch, 1974). At first sight, the recall data of the present experiment do not appear to fit this account, since they show a reliable effect of age and a reliable age × complexity interaction. However, the difference between the Baddeley and Hitch paradigm used by Morris et al. and the Daneman and Carpenter (1980) paradigm used in the present study is that the words to rehearse and recall are presented prior to the verification task in the former paradigm, whereas they must be added to the rehearsal loop during on-line processing in the latter. Results of the present study suggest that older people have greater difficulty with the ongoing processing aspects of working memory tasks, and thus they are less able to add additional words to the rehearsal loop, especially when complex sentences are presented.

At a less speculative level, the present study found no evidence for an age × pacing interaction in the present paradigm. There was also no evidence of an interaction between age and divided attention. Most interestingly, not all types of task complexity had similar effects, and this finding may prove useful both for dissecting the components of working memory and for understanding the particular processing difficulties associated with advancing age.

REFERENCES

BADDELEY, A. D., & HITCH, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-90). New York: Academic Press.

BADDELEY, A. D., LOGIE, R., BRESSI, S., DELLA SALA, S., & SPINNLER, H. (1986). Dementia and working memory. *Quarterly Journal of Experimental Psychology*, **38A**, 603-618.

CERELLA, J., POON, L., & WILLIAMS, D. (1980). Age and the complexity hypothesis. In L. Poon (Ed.), *Aging in the 1980s* (pp. 332-340). Washington, DC: American Psychological Association.

CHASE, W. G., & CLARK, H. H. (1972). Mental operations in the comparison of sentences and pictures. In L. W. Gregg (Ed.), *Cognition in learning and memory* (pp. 205-232). New York: Wiley.

COHEN, G. (1981). Inferential reasoning in old age. *Cognition*, **9**, 59-72.

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. 384-420). New York: Van Nostrand Reinhold.

CRAIK, F. I. M., & BYRD, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik & S. E. Trehub (Eds.), *Aging and cognitive processes* (pp. 191-211). New York: Plenum Press.

DANEMAN, M., & CARPENTER, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior, 19*, 450-466.

HITCH, G. J., & BADDELEY, A. D. (1976). Verbal reasoning and working memory. *Quarterly Journal of Experimental Psychology, 28*, 603-621.

INGLIS, J., & CAIRD, W. K. (1963). Age differences in successive responses to simultaneous stimulation. *Canadian Journal of Psychology, 17*, 98-105.

LIGHT, L. L., & ANDERSON, P. A. (1985). Working-memory capacity, age, and memory for discourse. *Journal of Gerontology, 40*, 737-747.

LIGHT, L. L., ZELINSKI, E. M., & MOORE, M. (1982). Adult age differences in reasoning from new information. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 8*, 435-447.

McDOWD, J. M., & CRAIK, F. I. M. (in press). The effects of aging and task difficulty on divided attention performance. *Journal of Experimental Psychology: Human Perception & Performance*.

MORRIS, R. G., GICK, M. L., & CRAIK, F. I. M. (1988). Processing resources and age differences in working memory. *Memory & Cognition, 16*, 362-366.

SALTHOUSE, T. A. (1982). *Adult cognition: An experimental psychology of human aging*. New York: Springer-Verlag.

SALTHOUSE, T. A., ROGAN, J. D., & PRILL, K. (1984). Division of attention: Age differences on a visually presented memory task. *Memory & Cognition, 12*, 613-620.

SPILICH, G. J. (1983). Life-span components of text processing: Structural and procedural differences. *Journal of Verbal Learning & Verbal Behavior, 22*, 231-244.

WELFORD, A. T. (1958). *Aging and human skill*. London: Oxford University Press.

WELFORD, A. T. (1977). Motor performance. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 450-496). New York: Van Nostrand Reinhold.

WRIGHT, R. (1981). Aging, divided attention, and processing capacity. *Journal of Gerontology, 36*, 605-614.

ZELINSKI, E. M., LIGHT, L. L., & GILEWSKI, M. J. (1984). Adult age differences in memory for prose: The question of sensitivity to passage structure. *Developmental Psychology, 20*, 1181-1192.

NOTE

1. It should be noted that this technique does not measure *memory span* in the strict sense. The measure is not the longest string recalled correctly, but simply the number of words correctly recalled. The terms *word span* and *sentence span* are retained, however, since with the small set sizes used, the present measures are quite comparable to traditional span measures.