British Journal of Psychology (1993), 84, 171–199 *Printed in Great Britain*

Influence of working memory on adult age differences in matrix reasoning

Timothy A. Salthouse*

School of Psychology, Georgia Institute of Technology, Atlanta, GA 30332-0170, USA

The four studies reported in this article, involving a total of 401 adults ranging between 18 and 80 years of age, were designed to investigate how working memory might mediate adult age differences in matrix reasoning tasks such as the Raven's Progressive Matrices Test. Evidence of this mediation is available in the finding that statistical control of an index of working memory reduces the age-related variance in matrix reasoning performance by approximately 70 per cent. Because the age differences were nearly constant across items of varying difficulty, it was concluded that the factors responsible for variation in item difficulty were distinct from those responsible for the age differences. However, young adults were found to be more accurate than older adults at recognizing information presented earlier in the matrix reasoning trial, thereby supporting the interpretation that working memory exerts its influence by contributing to the preservation of information during subsequent processing.

Speculations about the role of working memory in adult age differences in cognition can be traced at least as far back as Welford's (1958) book (see Salthouse, 1990, for a review of research on adult age differences in working memory). It has only been in the last several years, however, that convincing empirical evidence for the influence of working memory on age differences in cognition has become available. Examples are recent studies by Salthouse (1991*a*) and Salthouse, Mitchell, Skovronek & Babcock (1989) in which the magnitude of the age differences in several measures of cognitive functioning was found to be greatly attenuated by statistical control of a measure of working memory.

Although the statistical relations seem to be well documented, relatively little is yet known about the processes responsible for those relations. The principal goal of the current research was therefore to investigate the mechanisms by which working memory might mediate the age differences in a specific cognitive task. Three types of analyses are reported. The first focuses on item variation because items can be assumed to vary in the demands placed on working memory, and hence age × item interactions might be expected if high-demand items exceed the reduced working memory capacities of older adults, but are still within the capabilities of young adults. A second type of analysis is based on an examination of the alternatives selected on incorrectly answered items. Just as items can be postulated to vary in their working memory demands, so might the probability of selecting particular foils or incorrect alternatives differ according to the working memory requirements

---

* Requests for reprints.

imposed by each incorrect alternative and the working memory capabilities of the individual. Finally, because working memory is usually defined in terms of the simultaneous storage and processing of information, one should expect young adults, who tend to have high levels of working memory, to be better able to preserve previously presented information than older adults, who are presumed to have lower levels of working memory.

In addition to examining possible consequences of different levels of working memory, the results of these studies can also be used to investigate one possible cause of age-related variations in working memory. This is the idea that age differences in many cognitive tasks originate because of an age-related reduction in the speed of executing relevant cognitive operations. Some evidence for this interpretation was reported by Salthouse (1991*a*) and Salthouse & Babcock (1991) in which age differences in working memory were substantially reduced by using statistical procedures to equate people on a measure of perceptual comparison speed. The contribution of perceptual speed was examined in the current project by contrasting the degree of attenuation of the age differences in cognitive performance achieved by statistical control of a working memory measure with that obtained with control of a perceptual speed measure, and also by examining the additional attenuation attributable to working memory after the variance associated with perceptual speed had been removed.

Matrix reasoning tasks were selected as the criterion cognitive activity in this project because: (*a*) matrix reasoning tests such as the Raven's Progressive Matrices are a common and highly g-loaded (e.g. Jensen, 1982) measure of intelligence; (*b*) substantial age-related differences in Raven's Progressive Matrices performance have been reliably documented; (*c*) there are both theoretical and empirical reasons for expecting working memory to be important in the solution of these kinds of problems; (*d*) extensive development of the Raven's Progressive Matrices Test has resulted in a systematic progression of item difficulty from the early to the late items in the test; and (*e*) many of the incorrect alternatives in the Raven's problems were constructed to be informative about the causes of poor performance.

Research involving comparisons of adults of different ages on the Raven's Progressive Matrices Test has recently been reviewed by Salthouse (1992*c*). The median correlation between age and Raven's score across six studies involving samples with a wide range of ages was −.61. Five studies were located in which groups of older adults, typically with a mean age of about 70 years, were contrasted with groups of young adults, usually with a mean age of about 20 years. In each study, the performance of the older adults could be expressed in units of the distribution of the performance of young adults. The median across the five studies was −2.84 standard deviation units. Both of these estimates of the magnitude of the age differences are among the largest reported for any cognitive measure.

It has recently been reported that performance on the Raven's Progressive Matrices Test has increased over historical time (Flynn, 1987; Raven & Court, 1989), and some researchers have interpreted these positive time-lag differences as evidence that age differences observed in cross-sectional comparisons are an artifact of generational change. As discussed in Salthouse (1991*b*), a discovery of generational differences in test performance does not by itself imply that maturational factors do not contribute to the observed age-related differences because such a conclusion depends on several additional assumptions that are seldom tested. Regardless of the distal cause of performance differences among people of varying age, however, it is important to identify the proximal processing characteristics associated with different levels of performance. In other words, even if the age-related differences are an artifact of progressively higher test performance across successive generations, it still remains to be determined how individuals from earlier generations perform differently from individuals from later generations. From the current perspective, therefore, the goal is to specify the factors associated with differences in the observed level of functioning, and not to determine the ultimate source or cause of those differences.

The potential importance of working memory in successful performance of matrix reasoning problems can be illustrated by considering the sample problem in Fig. 1. Notice that all but one of the cells in the matrix of three rows and three columns is filled with geometric forms. The task for the examinee is to select which of the eight answer alternatives presented below the matrix provides the best completion of the missing cell of the matrix. One way to conceptualize the processing required in matrix reasoning problems such as this is in terms of the three components discussed by Jacobs & Vandeventer (1972): discrimination among elements, identification of relations and combination of relations. Working memory could be involved in each of these components because cell attributes have to be preserved in order to determine similarities and differences across cells, similarities and differences have to be preserved in order to determine the relations among cells in a given row or column, and relations have to be preserved in order to coordinate row and column relations to predict the pattern of the missing cell. Another theoretical analysis in which an aspect of working memory concerned with the management of subgoals was postulated to be critical for the solution of difficult matrix problems was also recently published by Carpenter, Just & Shell (1990). Finally, empirical support for the hypothesized relation between working memory and matrix reasoning performance is available in the moderate (.30 to .59) correlations reported by Larson and colleagues (e.g. Larson, Merritt & Williams, 1988; Larson & Saccuzzo, 1989) between Raven's performance and several measures postulated to reflect working memory abilities.

Evidence that items in the Raven's Test vary in item difficulty is available from Raven, Court & Raven (1985). These investigators reported that the average accuracy in the Advanced Set II version of the test was 98 per cent for item 1, 88 per cent for item 10, 59 per cent for item 20, and only 26 per cent for item 30. This substantial variation in item difficulty means that comparisons can be made at several points along the difficulty continuum. Moreover, to the extent that the item variation is at least partially attributable to differential demands on working memory, one might expect the effects associated with age to be greatest on the most difficult items. In other words, if one of the reasons for the low accuracy of difficult items is that those items place greater demands on working memory than less difficult items, then accuracy on those items should provide the greatest discrimination across people of different ages who are assumed to vary in their working memory abilities.

The sample problem in Fig. 1 can be used to illustrate how the particular incorrect

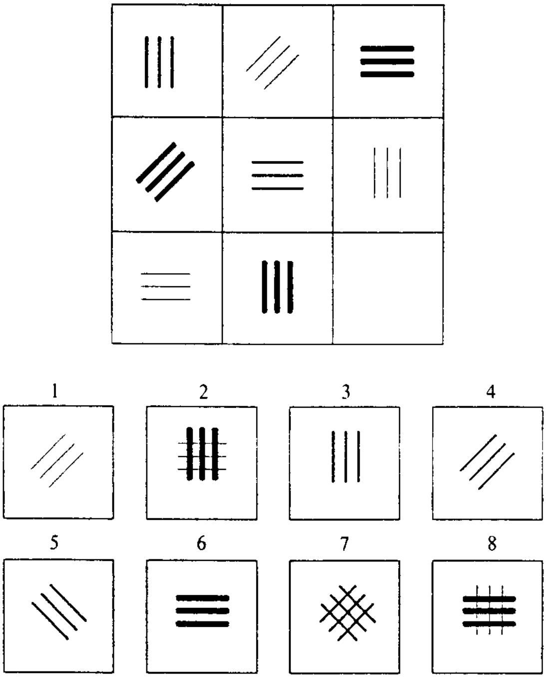


Figure 1. Examples of the type of problem in the Raven's Progressive Matrices Test.

choices selected might be informative about the causes for poor performance. Notice that the correct answer is alternative 4 because that pattern satisfies the critical principle of one of each type of line thickness and line orientation in each row and column. Alternative 6 is merely a repetition of a pattern from the matrix, and its selection may reflect a response based on the property of familiarity rather than a relevant principle. Alternatives 1 and 3 are also repetitions, but in addition can be considered partial solutions because they would have been correct if one dimension had not been neglected (i.e. line thickness in alternative 1, and line orientation in alternative 3). Alternative 5 is another example of a partially complete solution, in this case because the pattern does not incorporate the correct value for the relevant dimension of line orientation. Finally, alternatives 2, 7 and 8 are combinations created by applying an incorrect principle of addition. Some of these error types seem more likely to be a consequence of working memory limitations than others, and thus examination of the pattern of incorrect choices may prove informative about the influence of working memory on matrix reasoning performance. For example, a failure to notice a relevant attribute might occur equally often for people with effective and ineffective working memory systems, but people with more limited working memory systems might be expected to have a greater number of errors associated with a failure to coordinate the simultaneous application of multiple rules.

Results from four studies are reported in this article. Study 1 is a more detailed report of Study 1 from Salthouse (1991*a*), and focuses on the influence of working memory and age at the level of individual items in the Raven's Progressive Matrices

Test. The matrix reasoning task was implemented on a computer in Studies 2 through 4 in order to obtain separate measures of time and accuracy, and to allow for the presentation of probes of previously displayed information. Because recognition probes could be presented with either simultaneous or sequential displays of the stimulus matrix, and because both methods differ substantially from the more traditional paper-and-pencil format for presenting matrix reasoning problems, a primary purpose of Study 2 was to explore interrelations among measures of performance in the different presentation modes. Study 3 then examined accuracy of probe decisions after simultaneous displays of the matrix, and Study 4 examined probe decisions after sequential displays of the matrix. Participants in all studies also performed two perceptual speed tasks to allow comparisons of the influence of perceptual speed and working memory as possible mediators of age differences in matrix reasoning.

STUDY 1

The purpose of this initial study was to examine the influences of age and working memory on performance of the Raven's Progressive Matrices Test at the level of individual items. If the solution of certain items, or the selection of particular incorrect alternatives, depends on the individual's working memory ability, then the accuracy and error patterns should vary as a function of age and level of working memory.

Method

Subjects

Newspaper advertisements were used to recruit adults interested in participating in a research project concerned with relations among ageing, memory and cognition. Complete data were obtained from 221 adults (61 per cent women) ranging from 20 to 80 years of age. Each decade was represented by between 34 and 39 individuals. All participants reported the number of years of education they had received, and rated their health on a five-point scale (ranging from 1 for excellent to 5 for poor). Means of these variables and their correlations with age were 15.6 years and $r = -.04$ for education, and 2.04 and $r = .20$ ($p < .01$) for self-rated health. Analyses revealed that the patterns among the variables of primary interest were not significantly altered after adjusting for amount of education or rating of self-reported health, and consequently these measures are reported primarily to assist in description of the research sample.

Procedure

Research participants were tested in groups of about four to 30, and all performed the tasks in the same sequence. The tasks, in the order in which they were presented, were: Digit Symbol Substitution Test (Wechsler, 1981), Letter Comparison, Pattern Comparison, Computation Span, Listening Span Abstraction Test (Shipley, 1986) and Raven's Advanced Progressive Matrices, Set II (Raven, 1962). The Digit Symbol and Shipley Abstraction Tests were included for purposes unrelated to the current project and will not be discussed further in this report.

The Letter Comparison, Pattern Comparison, Computation Span and Listening Span tasks were identical to those described in Salthouse & Babcock (1991). The Letter Comparison and Pattern Comparison tasks consisted of pages containing pairs of three, six, or nine letters, or pairs of patterns composed of three, six or nine line segments. One-half of the pairs were identical, and one-half were different because of a change in one of the letters or line segments. The task for the participant was to classify each pair as SAME or DIFFERENT by writing an S or a D on a line between the two members of the pair as rapidly as possible. Trials with three, six or nine elements were separately timed (30 s for

32 pairs), and the scores were averaged to provide a single measure for each type of comparison (letter or pattern).

The Computation Span and Listening Span tasks were designed to assess working memory by requiring participants to remember information while also carrying out specified processing. In the Computation Span task arithmetic problems were presented auditorily, and the task was to select the answer to the arithmetic problem from three alternatives on the response form while also remembering the last digit in each problem. Short sentences were auditorily presented in the Listening Span task, with participants instructed to answer a question about the sentence, printed on the response form along with three answer alternatives, while also remembering the last word in each sentence. Target items were recalled by writing them, in the order in which they were presented, on designated lines on the back of the response form. The number of arithmetic problems or sentences increased from one to seven, with three trials at each sequence length. An individual's span was determined by the greatest number of digits or words that could be remembered on two of the three trials for a particular sequence length, given that he or she was also correct in the answers to the relevant arithmetic and sentence comprehension questions. This latter requirement ensured that the scores represented both storage and processing.

As noted above, the Raven's Advanced Progressive Matrices Test consists of displays of 3 × 3 matrices of geometric forms with the bottom right cell missing. The task for the examinee is to select the correct pattern to complete the missing cell from a set of eight alternatives displayed below the matrix. Three sample problems (Items 6, 8 and 12 from the Raven's Advanced Progressive Matrices Set I) were provided, followed by the 36 problems of Set II. Twenty minutes were allowed to work the problems in Set II. Although a time limit of 40 min is usually recommended, the steep difficulty gradient across successive items suggests that few responses were likely to have been correct on later items had they been attempted. Furthermore, comparisons of the average number of items answered correctly in college student samples with 20 min (a pilot study, and Studies 2 and 4 of the present report), 40 min (Larson & Saccuzzo, 1989; Palmer, MacLeod, Hunt & Davidson, 1985) and no time-limit (Jensen, 1983, 1987; Paul, 1985) administrations revealed relatively small differences related to the time allowed for completion of the items. Finally, similar age trends were reported by Heron & Chown (1967) with 20 min and 40 min administrations of the standard Raven's Progressive Matrices Test.

Results and discussion

Distributions of the frequency of number-correct scores on the Raven's Progressive Matrices Test for each of three age groups ($N = 77$, 73 and 71, respectively) are displayed in Fig. 2. Also portrayed in Fig. 2 for purposes of comparison is the distribution of scores obtained in a pilot sample of 83 college students (mean age = 19.9 years, mean of 13.7 years of education). It is apparent in this figure that college students have the highest scores, and that increased age is associated with a systematic downward shift of the entire distribution.[1]

The initial step in the analyses of the data involved creating composite measures of the working memory and perceptual speed constructs. This was accomplished by averaging the individual's z scores from the Listening Span and Computation Span tasks to produce a working memory composite, and averaging his or her z scores from the Letter Comparison and Pattern Comparison tasks to produce a perceptual speed composite.

Table 1 contains the correlation matrix summarizing the relations among the variables of subject age, the two composite variables, and both the number of correct

---

[1] The higher scores for the sample of students relative to the sample of adults between 20 and 39 years of age may be a consequence of their younger age (i.e. mean of 19.9 years versus 29.1 years), or of the fact that they had been selected partially on the basis of intellectual ability for admission to a relatively competitive university. If the latter is the case then comparisons between students and older adults recruited according to the procedures in this study may overestimate the absolute magnitude of age-related differences.

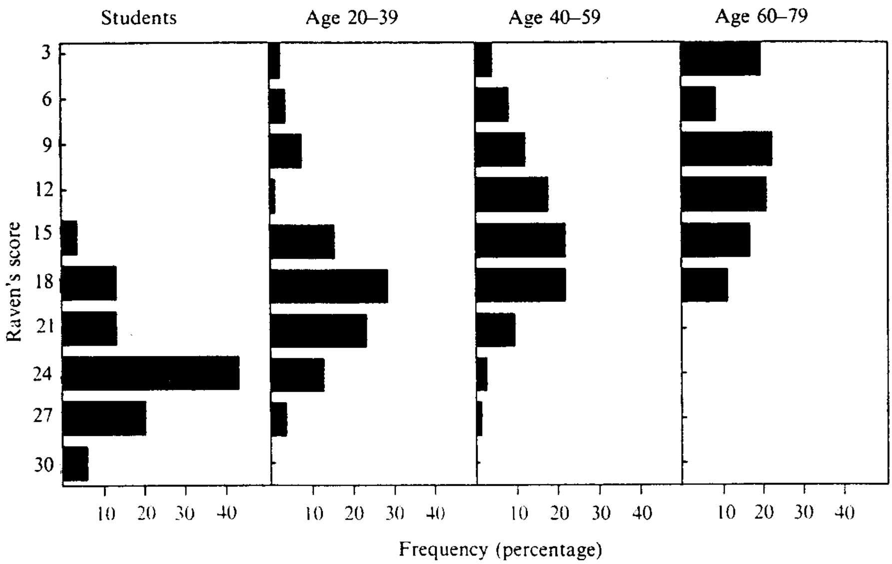


Figure 2. Distribution of scores on the Raven's Progressive Matrices Test as a function of age, Study 1.

responses and the percentage of attempted items answered correctly in the Raven's Progressive Matrices Test. It should be noted that each measure had at least moderate reliability, and that all correlations were in the medium to large range. Moreover, the significant negative correlation between age and percentage of items answered correctly indicates that not all the age variation in Raven's performance is due to differences in the number of items attempted.

Table 1. Correlation matrix for variables in Study 1 (*N* = 221)

Variable	1	2	3	4	5
1 Age	×	−.57*	−.36*	−.54*	−.61*
2 RPM no.C		(.91)	.84*	.69*	.62*
3 RPM%C			(.86)	.60*	.45*
4 WMem				(.71)	.59*
5 Speed					(.85)
Mean	48.5	14.0	63.3	0.00	0.00
SD	17.4	6.2	23.7	0.88	0.93

* $p < .01$.

Note. Reliabilities in the diagonals were estimated by using the Spearman–Brown formula to boost the correlation between the two component measures (for Perceptual Speed and Working Memory) or between the scores for odd and even items (for the number-correct and percentage-correct variables).
Key. Speed = Average of z scores for Letter Comparison and Pattern Comparison; WMem = Average of z scores for Computation Span and Listening Span; Rav no. C = Number of items correct in Raven's Test; Rav%C = Percentage of attempted items correct in Raven's Test.

A series of hierarchical multiple regression analyses was conducted to determine the amount of variance in measures of Raven's performance associated with age when considered alone, and after controlling the variance associated with the composite measure of working memory and/or the composite measure of perceptual speed. The R^2 values in these analyses are summarized in the top row of Table 2, where it can be seen that partialling the variance associated with working memory reduced the R^2 for age from .322 to .053. This degree of attenuation of the age effects is clearly consistent with a mediational influence of working memory on the age differences in matrix reasoning. It is also apparent in Table 2 that the attenuation of the age differences was nearly as large after statistical control of the perceptual speed variable. Implications of this finding will be considered in the General Discussion.

Table 2. R^2 estimates associated with age in prediction of matrix reasoning performance

Study	Variable	Alone	After WMem	After PSpeed	After WMem & PSpeed
All subjects					
1	Raven's Num. Correct	.322*	.053*	.056*	.017*
2	Raven's Num. Correct	.678*	.389*	.201*	.162*
	Simult. Matrix Acc.	.347*	.108*	.062	.031
	Sequent. Matrix Acc.	.228*	.066	.021	.007
3	Simult. Matrix Acc.	.375*	.086*	.111*	.021
Only subjects satisfying accuracy criterion					
1	Raven's Num. Correct	.303*	.109*	.083*	.048*
2	Raven's Num. Correct	.701*	.482*	.134*	.117*
	Simult. Matrix Acc.	.258*	.157*	.111	.107
	Sequent. Matrix Acc.	.138	.049	.005	.002
3	Simult. Matrix Acc.	.273*	.137	.058	.039

* $p < .01$.

In an attempt to minimize the possibility that the relations between working memory and Raven's performance might have been a consequence of a failure on the part of some subjects to understand fully the task requirements, the analyses were repeated after eliminating subjects with errors on either of the first two items in the Raven's Set II problems. Results from this restricted sample of 166 subjects are summarized in the bottom portion of Table 2. Notice that the pattern is qualitatively very similar to that evident in the complete sample.

Additional analyses were conducted to examine performance on the Raven's Test at the level of individual items. These analyses were limited to items 1 through 22 because fewer than 50 per cent of the research participants responded to later items, and the range of ages for those who did respond was greatly restricted because few older individuals attempted many of the later items.

The percentages of correct responses as a function of item across the three age groups are displayed in Fig. 3. The most striking feature of these data is that although

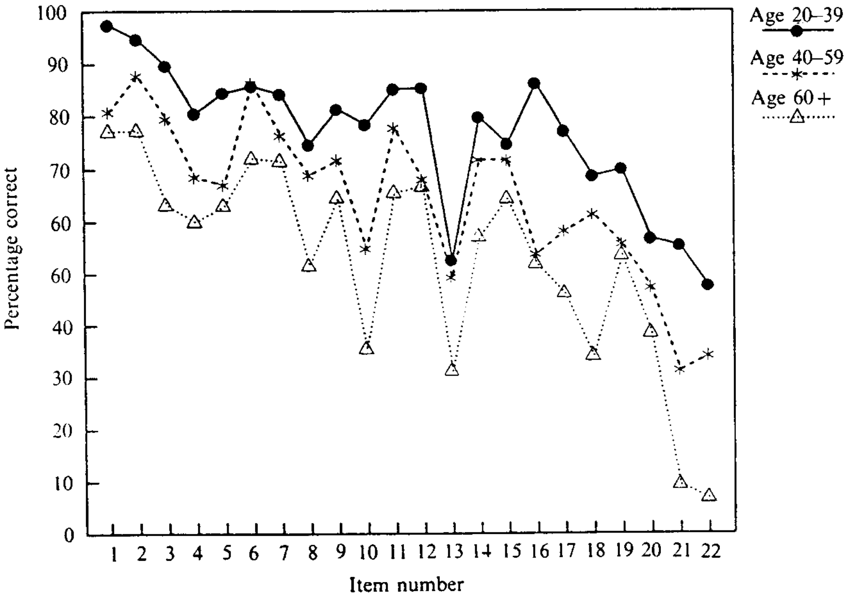


Figure 3. Mean accuracy for individual items in the Raven's Progressive Matrices Test as a function of age, Study 1.

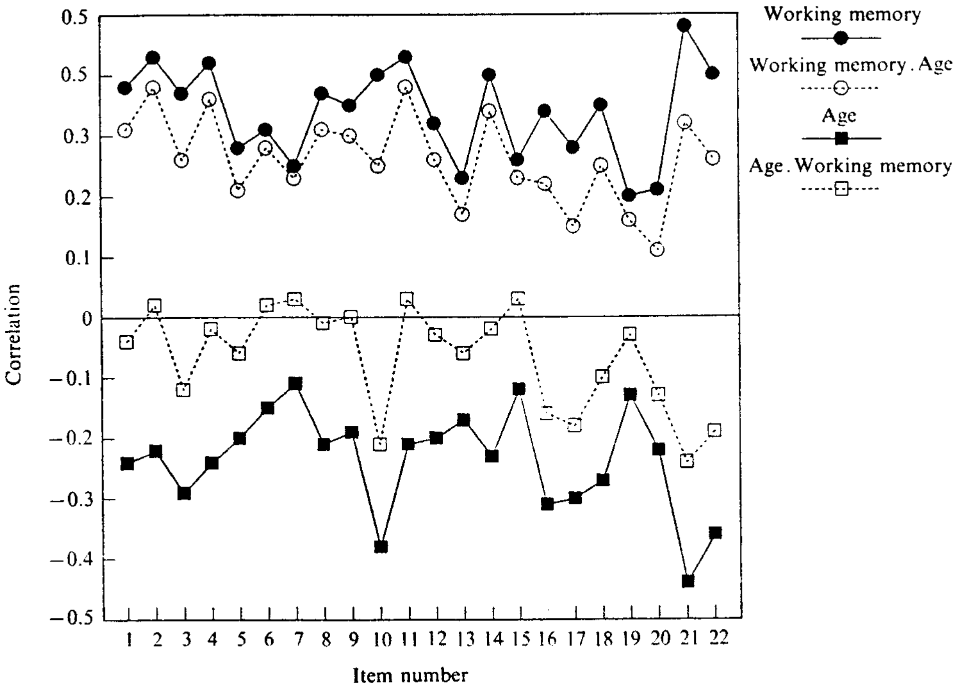


Figure 4. Correlations of age and working memory with accuracy of individual items on the Raven's Progressive Matrices Test, Study 1.

Table 3. First and second most frequently chosen incorrect alternatives for each item by age group

	Number of errors			1st frequent			2nd frequent		
Item no.	Y	M	O	Y	M	O	Y	M	O
1	3	13	15	1	1,7	1	—	1,7	6,7,8
2	4	9	15	2,4	2,4,6	2	2,4	2,4,6	6
3	9	14	25	8	8	8	4	1,5,6	3
4	17	20	26	8	8	5	7	2	8
5	13	23	24	1	1,7	7	2	1,7	1
6	11	10	19	7	7	7	2	6,8	2
7	14	14	21	2	2	2	5	5	4
8	21	19	32	2	2	3	8	3,5	5
9	15	20	23	1	1	1	7	3	7
10	18	30	39	8	8	8	2	2	7
11	13	14	21	7	7	7	4	4,6	1,3,4
12	13	21	21	5	5	5	8	3,8	8
13	37	37	43	5	5	5	6	6,7	6
14	16	19	25	4	4	4,7	7	7	4,7
15	21	18	19	8	8	6	6	1,4	4,7
16	11	29	27	5	5	5	6,7	1	6
17	17	27	29	3	3	3	4	4	4
18	25	20	27	1	1	1	3	5	5
19	23	25	19	5	5	5	8	6	1,6,7,8
20	30	27	24	2	2	2	4	4	4
21	31	30	29	1	1	1	4	4	4
22	37	24	27	8	8	8	2,5	4	1

there is substantial variation in average level of accuracy, ranging from about 85 per cent on item 1 to approximately 35 per cent on item 22, the functions for the three groups appear remarkably parallel. This impression was reinforced with the results of a group (young, middle, old) × item (items 1 to 22) analysis of variance. Both main effects of age group ($F(2,218) = 30.94$, MSE = 1.07, and item, $F(21,4578) = 47.69$, MSE = 0.15) were significant ($p < .01$) in this analysis, but their interaction was not ($F(42,4578) = 1.41$).

A second analysis consisted of determining the simple and partial correlations between accuracy for each item and the age and working memory variables. These correlations, displayed in Fig. 4, serve to confirm the inference from Fig. 3 of nearly constant age effects across items. It is also interesting that partialling working memory from the age correlations reduced most of them to near zero, but that the reduction in working memory correlations after partialling age was much smaller. Medians of the 22 correlations were −.22 for age alone, −.04 for age after partialling working memory, .35 for working memory alone, and .26 for working memory after partialling age.

Patterns of errors on items with overt responses (i.e. excluding errors of omission) were also examined. Analyses were conducted to determine whether subjects of varying ages differed in the particular incorrect alternatives selected with the greatest frequency. Table 3 contains the identities of the alternatives for each item with the highest, and second highest, frequency of occurrence in each age group. It is apparent that there was a great deal of consistency in the particular incorrect alternatives chosen by adults of different ages. The young and middle-aged groups selected the same incorrect alternative most frequently on all 22 items, and there was agreement between the young and old groups, and between the middle-aged and old groups, on 19 of the 22 items. Furthermore, there was agreement in the alternatives selected with either the highest or the second highest frequency in 21 of the 22 items both for middle-aged and old groups, and for young and old groups.

Discussion

The results summarized in Tables 1 and 2 confirm previous results concerning relations among age, working memory, and measures of cognitive performance. Specifically, the finding that the influence of age on cognitive performance was markedly reduced after controlling the variance associated with a measure of working memory suggests that working memory plays an important role in the age difference in at least this cognitive task.

Average solution accuracy varied considerably across the items examined, and it seems reasonable to hypothesize that at least some of the item variation might have been due to increased working memory demands. If age differences are largely mediated by reductions in working memory, then it might have been expected that the differences would be greater for the most difficult items that place the largest demands on working memory. Contrary to this prediction, however, the analyses revealed that the age relations were remarkably constant across the range of items examined. An apparent implication of these findings is that the factors responsible for the variation in item difficulty are independent of the factors responsible for the effects associated with working memory and adult age. Of course, it is not known whether similar results would have been observed in samples with higher average levels of reasoning performance, of if it had been possible to extend the analyses to even more difficult items, but the patterns in Figs 2 and 3 suggest that the influences of age and working memory are relatively constant in this set of data.

The analyses of the error patterns also failed to provide any evidence of deficits in particular kinds of processing, or restricted to specific relational principles. Although some incorrect alternatives were consistently selected more often than others, the same general patterns were evident across the three age groups. It therefore appears that the factors responsible for causing some incorrect alternatives to be chosen more frequently than others were operating in similar ways regardless of the individual's age.

The configuration of results just described presents a challenge for interpretation. On the one hand, there is evidence of moderate to large relations between the measures of working memory and matrix reasoning performance, but on the other hand, the data indicate that these relations are no greater for difficult (low accuracy) problems than for easy (high accuracy) problems. At least three distinct explanations could be proposed to account for these findings. First, it is possible that the lack of a differential relation across items of varying difficulty is a consequence of some type of methodological artifact. For example, the fact that the most difficult items occurred later in the problem sequence means that fewer of them were probably attempted by the slower subjects, who may also have had less effective working memory systems. An attempt was made to minimize distortions of this kind by restricting the analyses to items attempted by the majority of the subjects, but it is nevertheless still possible that various kinds of measurement insensitivity may have contributed to the lack of differential relations of age and working memory across items.

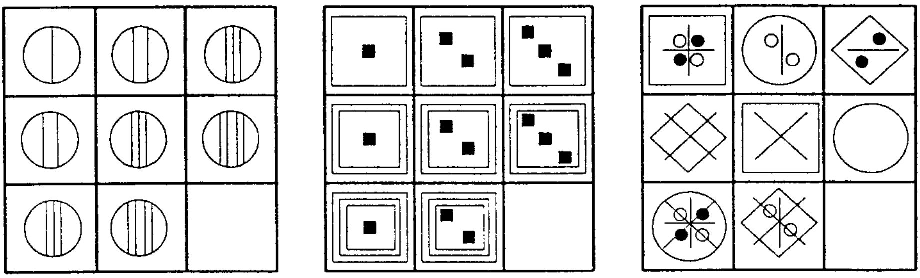
A second potential explanation is that much of the variation in item difficulty may be attributable to factors unrelated to working memory. That is, the working memory influence might have been nearly the same for all items, but item difficulty could have varied because of factors such as salience of the attributes, familiarity of the relational rules, etc.

A third interpretation focuses on the meaning of the relation observed between the working memory measures and performance on the matrix reasoning test. That is, it could be speculated that this relation did not originate because of the involvement of working memory in the reasoning test, but rather was a consequence of a third variable involved in both the working memory and reasoning tests. As an example, instead of the relation occurring because of reliance on working memory in the matrix task, it could have originated because of a temporal aspect in both the matrix reasoning test and in the working memory tests. Consistent with this view is the fact that the 20 min time limit for the Raven's Test was too short for most respondents to attempt all items, and the group administration of the working memory tests necessitated the imposition of limits on the time between presentation of successive stimuli and the time allowed for production of responses. It is therefore conceivable that at least some of the relations observed in this study were a reflection of the common requirements of having to work rapidly, and were not a direct consequence of working memory involvement in the Raven's Matrix Test.

STUDY 2

The remaining studies in this project were designed in part to resolve some of the interpretational ambiguities associated with the results of Study 1. Features of these studies intended to address the concerns discussed above were: (*a*) both the working memory and matrix reasoning tests were administered on a computer with subjects allowed as much time as desired to respond; and (*b*) matrix problems created to vary systematically according to the number of relations among matrix elements were presented in a randomly arranged sequence.

Examples of the problems used in Studies 2 and 3 are illustrated in Fig. 5. Notice that the problems vary with respect to the number of unique relations among cell attributes. That is, cells in the problem on the left vary only in the number of vertical lines, cells in the middle problem vary in the number of filled squares and in the number of surrounding squares, and cells in the problem on the right vary in the



One relation Two relations Three relations

Figure 5. Sample problems with one, two and three relations.

distribution of major shape (circle, square and diamond), in the distribution of inner patterns (filled and open circles), and in the distribution of inner lines (vertical, horizontal and diagonal).

If the problems with more relations impose greater demands on working memory than problems with fewer relations, then one might expect progressively higher correlations between working memory and decision accuracy as the number of relations in the problem increased. Furthermore, if adult age differences in matrix reasoning are attributable to limitations of working memory, then the magnitude of the age differences should also increase as a function of the number of relations in the problem.

Another purpose of Studies 2, 3 and 4 was to examine the implication that people with higher scores on the working memory measures should be more accurate at recognizing previously presented information than people with lower working memory scores. Although the information probe procedure is conceptually very simple, it can become complicated in practice because of the difficulty of specifying the appropriate type of previously processed information to be probed. For example, the most relevant processing in matrix reasoning tasks may be intermediate between the initial encoding of the cell attributes and determination of the pattern for the missing cell. Unfortunately, probes concerning the status of information about similarities or differences across cells, row or column relations, etc., would be difficult to interpret without some assurance that the probed information had, in fact, existed in memory at an earlier time. That is, because some subjects may not generate these intermediate information states, and because people might vary in how the intermediate information is represented memorially, there can be formidable difficulties associated with probing the status of information assumed to exist in working memory at an earlier time. The solution to this problem adopted in the current studies involved administering probes about previously presented information—that is, probes of the patterns presented in specific cells of the matrix.

This approach to probing earlier information is not optimal because it is based on the unverified assumption that explicitly presented information is retained in working memory in a relatively untransformed state. It may be equally plausible to suggest that information in working memory is always transformed in some fashion, and does not merely duplicate the information available in the external stimulus. Although these concerns are legitimate, certain assumptions are necessary if the status of previous information is to be probed. Moreover, the validity of the assumption concerning the existence of previously presented information in working memory may depend on how, and when, the stimulus displays and information probes are presented. Consider the two presentation methods illustrated in Fig. 6.

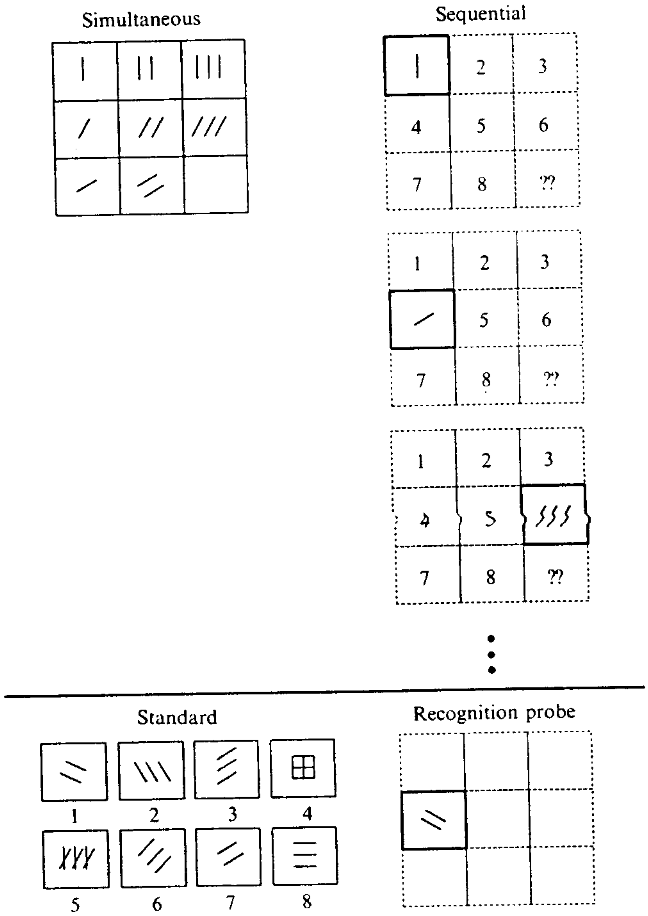


Figure 6. Illustration of the simultaneous and sequential methods of matrix presentation, and the standard and recognition probe displays in the computer-administered matrix reasoning task.

The simultaneous version of the matrix task portrayed in the left of Fig. 6 involves the presentation of the stimulus matrix until the subject indicates that he or she is ready to view the answer alternatives, and then substituting a recognition probe for the answer alternatives on some of the trials. A potential problem with this procedure is that even if the untransformed cell information had previously existed in working memory, it may have been discarded by the time the subject had completed processing the matrix and was ready to view the answer alternatives. This problem might be avoided with the sequential version of the task portrayed in the right of Fig. 6 because this mode of presentation allows the subject's processing to be interrupted by a probe of a previously viewed cell. Information regarding the contents of previously viewed cells may be more likely to still reside in working memory if the probe is presented while the subject remains engaged in processing the stimulus matrix.

Before investigating possible age-related differences in probe recognition accuracy it was first considered desirable to examine relations among performance in the Raven's Progressive Matrices Test and the two computer-administered matrix tasks illustrated in Fig. 6 when they were presented without any probe trials. Study 2 was therefore conducted to examine the relations among the matrix reasoning measures in the three presentation modes. The simultaneous task was then administered with recognition probes in Study 3, and the sequential task with recognition probes was administered in Study 4.

Method

Subjects

Descriptive characteristics of the 30 young adults and 30 older adults who participated in this study are summarized in Table 4. The young adults were college students, and the older adults were volunteers recruited from newspaper advertisements and referrals from other participants. Years of education refers to the number of years of formal education the participants reported having completed. Self-rated health is a subjective evaluation of one's health on a five-point scale ranging from 1 for excellent to 5 for poor. Digit–symbol time is the median time in s per digit–symbol pair, and digit–digit time is the median time in s per digit–digit pair.

Procedure

All subjects performed the tasks in the same fixed order: Raven's Advanced Progressive Matrices; Digit Symbol; Digit Digit; Reading Span; Computation Span; Simultaneous Matrix; and Sequential Matrix. Each task was preceded by instructions and several practice trials to attempt to ensure that the subjects clearly understood what they were supposed to be doing.

The Raven's Advanced Progressive Matrices Test was administered in the same manner (i.e. with the same time limits and the same practice problems) as in Study 1. The Digit–Symbol Test was a computer-administered version of the Digit–Symbol Substitution Test (Wechsler, 1981). The code table with the nine digit–symbol pairs was presented on the top of the computer screen, and a single digit–symbol pair was presented in the middle of the screen. On half of the trials the digit–symbol pair was correct, in that the pair matched that presented in the code table, and on half of the trials the digit–symbol pair was incorrect. Five trials of each type were presented with each of the nine digits (1 through 9). Subjects were instructed to classify the pair as CORRECT or INCORRECT as rapidly as possible by pressing the '/' key for correct and the 'Z' key for incorrect.

The Digit–Digit Test was similar to the Digit–Symbol Test except that the code table consisted of pairs of identical digits, and the test stimulus consisted of a pair of digits. On half of the trials the two digits in the pair were identical, and on half of the trials the digits were different. Five trials of each type were presented with each of the nine digits as the top member of the pair. Subjects were instructed to classify the pair as CORRECT or INCORRECT as rapidly as possible by pressing the '/' key for same or correct, and the 'Z' key for different or incorrect.

The Reading Span and Computation Span tasks were adaptations, with the same stimulus materials, of the two working memory tasks used in Study 1. Each required the research participant to demonstrate successful processing while also remembering information. The Reading Span task consisted of displays of sentences, each accompanied by a short question and three alternative answers. Subjects were instructed to read the sentence, type in the number designating the correct answer to the question, and remember the last word of the sentence. The Computation Span task consisted of displays of arithmetic problems with three alternative answers. Subjects were instructed to solve the arithmetic problem, type in the number designating the correct answer to the question, and remember the last digit in the problem. No limit was imposed on the duration each sentence or arithmetic problem could

Table 4. Characteristics of the samples in Studies 2 through 4

	Study 2	Study 3	Study 4
Sample size			
Young	30	30	30
Old	30	30	30
% Female			
Young	50	50	50
Old	53	50	57
Age			
Young	20.3 (1.5)	19.8 (1.1)	19.7 (1.7)
Old	61.5 (7.9)	60.8 (5.3)	63.4 (7.6)
Years of education			
Young	14.1 (1.3)	13.9 (1.0)	13.5 (1.4)
Old	15.1 (2.3)	15.4 (2.3)	15.0 (2.2)
Self-rated health			
Young	1.6 (0.7)	1.4 (0.6)	1.5 (0.6)
Old	1.7 (0.9)	1.8 (0.7)	1.9 (1.00)
Digit-Symbol Time (s)			
Young	1.20 (0.17)	1.17 (0.19)	1.13 (0.22)
Old	1.91 (0.41)	1.97 (0.36)	1.63 (0.28)
Digit–Digit Time (s)			
Young	0.54 (0.06)	0.55 (0.07)	—
Old	0.78 (0.26)	0.79 (0.19)	—
Computation Span			
Young	4.9 (1.9)	4.7 (2.0)	—
Old	3.1 (2.4)	2.1 (2.0)	—
Reading Span			
Young	3.2 (1.6)	3.4 (1.6)	—
Old	2.2 (1.6)	1.8 (1.3)	—
Raven Progressive Matrices Number Correct			
Young	23.4 (2.9)	—	20.5 (3.3)
Old	13.1 (5.4)	—	11.9 (5.9)

Note. Values in parentheses are standard deviations.

be viewed, but subjects were encouraged not to spend too much time on any given display. Pressing the ENTER key after viewing the sentence or arithmetic display resulted in the presentation of the next item in the sequence, or the word RECALL accompanied by a blank line for each to-be-recalled item. Spans were determined by increasing the number of sentences or arithmetic problems presented before the recall instruction. Three trials were presented with each sequence length, in an ascending order, until the subject was incorrect on either the processing or the recall on two of the three trials for a given sequence length. The subject's span was therefore the largest number of words or digits that could be accurately recalled on at least two of three trials while also correctly performing the required processing.

Two sets of 30 problems each were constructed for the computer-administered matrix tasks. Within each set, 12 of the problems had one relevant relation, 12 had two relevant relations, and six had three relevant relations. The assignment of problem set to task version (simultaneous or sequential) was balanced across subjects in each age group. (Preliminary analyses revealed no main effects or interactions of stimulus set and hence this variable is ignored in subsequent analyses.) The seven incorrect answer alternatives for each problem were constructed to represent a range of likely errors, such as failure to notice a relevant attribute, application of an incorrect relation, etc.

Instructions in both versions of the computer-administered matrix tasks emphasized that the tasks were very similar to the paper-and-pencil Raven's Progressive Matrices Test performed earlier except that the problems were now presented on the computer. Subjects were allowed to devote as much time as desired to inspecting the first display in the simultaneous version of the task, but they were not allowed to return to the matrix after viewing the answer alternatives. In the sequential version of the task the subjects were encouraged to keep the total number of matrix cells examined per trial as low

Table 5. Mean inspection time, decision time, and accuracy in the simultaneous and sequential matrix tasks, Studies 2, 3 and 4

	Simultaneous		Sequential	
	Mean	SD	Mean	SD
	Inspection time (s)			
Study 2				
Young	12.24	3.35	13.97	5.75
Old	27.62	13.14	39.71	19.03
t(58)	6.21*		7.09*	
Study 3				
Young	14.68	2.88	—	—
Old	29.65	10.49	—	—
t(58)	7.54*			
Study 4				
Young	—	—	15.39	6.01
Old	—	—	27.97	11.89
t(58)			5.18*	
	Decision time (s)			
Study 2				
Young	2.95	0.57	2.92	0.75
Old	5.03	1.66	6.30	3.02
t(58)	6.49*		5.95*	
Study 3				
Young	3.50	0.65	—	—
Old	6.55	2.14	—	—
t(58)	7.47*			
Study 4				
Young	—	—	2.54	0.39
Old	—	—	6.39	5.33
t(58)			3.94*	
	Decision accuracy (percentage correct)			
Study 2				
Young	84.6	9.5	70.7	15.4
Old	63.0	21.2	52.3	21.7
t(58)	5.08*		3.80*	
Study 3				
Young	74.6	12.4	—	—
Old	54.7	15.6	—	—
t(58)	5.47*			
Study 4				
Young	—	—	89.6	7.2
Old	—	—	54.5	25.3
t(58)			7.32*	

* $p < .01$.

as possible, but they were allowed to inspect the pattern in any cell as many times, and for as long a duration, as desired. As in the simultaneous version, subjects were not allowed to return to the matrix display after viewing the answer alternatives.

Results and discussion

Descriptive statistics for several of the performance measures are summarized in Table 4. The age differences in the time and span measures were statistically significant ($p < .01$), as was the difference on the Raven's Progressive Matrices measure (all ts > 3.0).

Means and standard deviations for the time and accuracy measures from the computer matrix tasks are displayed in Table 5. Inspection time refers to the time spent examining the stimulus matrix, and decision time refers to the interval from the display of the answer alternatives until the response indicating the subject's decision. It is apparent from these data that older adults took significantly more time than young adults in both the inspection and decision phases of the trial. However, it is important to note that the decisions resulting from this greater time were still significantly less accurate than those of young adults.

The correlation matrix containing the correlations among the major variables is presented in Table 6. Note that the correlations among the measures of matrix reasoning performance, and the correlations of these measures with the composite working memory measure, are all statistically significant, and moderate to large in magnitude. The first result suggests that common processes are probably involved in the three matrix tasks despite the quite different presentation formats. The second result confirms the finding of Study 1 that higher working memory scores are associated with more successful matrix reasoning performance.

Several hierarchical regression analyses were conducted in order to examine the relations among age, working memory and matrix reasoning performance. For each of the versions of the matrix task, the R^2 associated with age was determined when considered alone and after statistical control of the composite measures of working memory and perceptual speed separately, and in combination. Results of these analyses are summarized in the second, third and fourth rows of Table 2, along with the results of a parallel set of analyses based on data from subjects for whom one can have some confidence that they clearly understood the task requirements (corresponding rows of bottom panel). Evidence of this understanding was manifested by errorless performance on problems 1 and 2 in the Raven's Progressive Matrices Test, or by accuracy equal or greater than 75 per cent on the simplest one-relation problems in the simultaneous and sequential matrix tests. The numbers of young and old adults meeting these criteria were, respectively, 29 and 23 for the Raven's Test, 29 and 16 for the simultaneous matrix test, and 17 and 12 for the sequential matrix test. The major result apparent in the relevant rows of Table 2 is that all three measures of matrix reasoning performance exhibited the same pattern of attenuated age differences after statistical control of working memory. The larger age effects on the Raven's Test in this study relative to Study 1 are probably a consequence of the extreme groups design, but the important point is that the pattern of attenuation of the age differences was very similar. Furthermore, as was the case in Study 1, statistical control of the perceptual speed variable resulted in at least as

Table 6. Correlation matrix for variables in Study 2 ($N = 60$)

Variable	1	2	3	4	5	6	7
1 Age	×	−.82*	−.59*	−.48*	.44*	−.46*	.76*
2 RPM		(.94)	.80*	.72*	−.35*	.59*	−.70*
3 SimAcc			(.88)	.73*	−.17	.64*	−.56*
4 SeqAcc				(.84)	−.08	.54*	−.51*
5 Repet.					×	−.26	.29*
6 WMem						(.63)	−.47*
7 Speed							(.79)
Mean	40.9	18.2	73.8	61.5	8.0	0.00	0.00
SD	21.5	6.7	19.6	20.9	6.2	0.86	0.91

* $p < .01$.

Note. Reliabilities in the diagonals were estimated by using the Spearman–Brown formula to boost the correlation between the two component measures (for Perceptual Speed and Working Memory) or between the scores for odd and even items (for the matrix reasoning measures).

Key. RPM = Number correct in the Raven Progressive Matrices Test; SimAcc = Accuracy in the simultaneous condition of the computer matrices task; SeqAcc = Accuracy in the sequential condition of the computer matrices task; Repet. = Average number of repetitive cell examinations in the sequential condition of the computer matrices task; WMem = Working memory composite created by averaging *z* scores from the Reading Span and Computation Span tasks; Speed = Speed composite created by averaging *z* scores from the Digit–Symbol and Digit–Digit tasks.

much attenuation of the age differences in matrix reasoning as did the working memory variable.

Decision accuracy was also analysed according to the number of attribute relations in the problem. Mean accuracy in the simultaneous condition is displayed in the top panel of Fig. 7, and mean accuracy in the successive condition is displayed in the bottom panel of this figure. The lack of differential age effects apparent in the figure was confirmed by the absence of significant (i.e. $F < 1$) age × problem type × condition interactions in an age (young, old) × condition (simultaneous, sequential) × problem type (1, 2 or 3 relations) analysis of variance.

Correlations were also computed between working memory and accuracy for problems with each number of relations. The correlations (all significant at $p < .01$) for problems with one, two and three relations were .56, .60 and .51 for the simultaneous condition, and .46, .54 and .40 in the sequential condition. These results are consistent with the nearly constant working memory correlations across items evident in Fig. 4.

Because subjects in the sequential condition made overt keypress responses to examine the contents of each matrix cell, the pattern of cell examinations could also be analysed. Of the two types of cell examinations—unique or first examinations of a cell, and repetitive or redundant cell examinations—the latter are of greatest interest in this context. It can be seen in Table 6 that there was a significant correlation between age and number of repetitive cell examinations. It is tempting to infer that the greater number of redundant cell examinations on the part of older

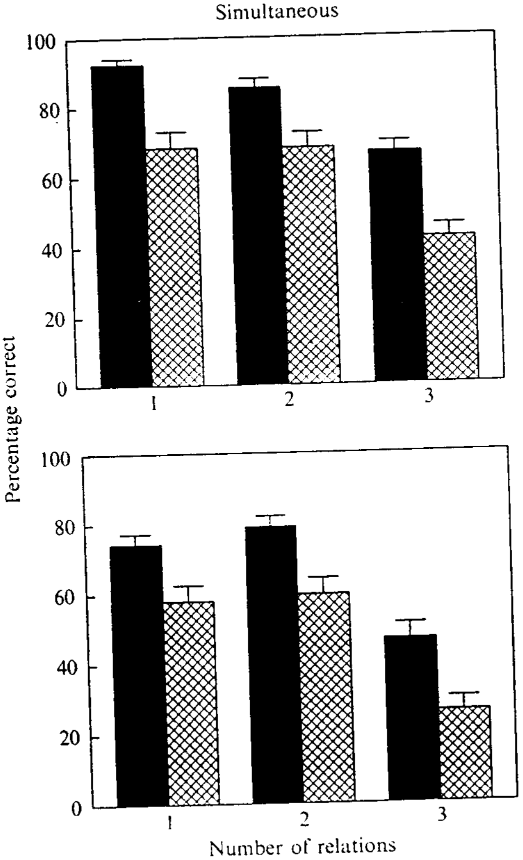


Figure 7. Mean accuracy for young and old adults as a function of the number of relations in the problem, Study 2. Lines above each bar correspond to one standard error. ■, young; ☒, old.

adults reflects an attempt to compensate for a diminished ability to preserve information in working memory. However, the low correlation between the number of repetitive cell examinations and the composite working memory measure (i.e. −.26, and only −.12 for young adults and −.09 for older adults) is inconsistent with this interpretation. Alternative causes of the repeated cell examinations may be difficulties in higher-order processing such as verifying hypotheses about row or column relations, or lack of confidence about the status of information in one's memory.

There appear to be three major findings from this study. The first is that the earlier results of moderately large relations among age, working memory and matrix reasoning performance were confirmed in measures without external time limits. These relations therefore cannot be attributed to a common temporal limitation imposed by the group-administration procedures in the working memory and matrix reasoning tasks. A second major finding is the replication of the results of Study 1 that the age and working memory influences were nearly constant across problems of varying difficulty (and presumably, working memory requirements). Finally, the discovery that the age and working memory influences were similar across the three matrix tasks, and that the measures of performance were significantly correlated with one another, suggest that common processes were involved in each task despite the quite different methods of presentation.

STUDY 3

The primary purpose of Study 3 was to investigate the interrelations among age, measures of the accuracy of decisions concerning probed information, accuracy on the simultaneous matrix task, and performance on the working memory tasks. If retention of earlier information is important for successful performance on matrix reasoning tests, then one would expect moderate positive correlations between probe accuracy and matrix performance. Furthermore, if limitations of working memory contribute to poor matrix performance because of an inability to preserve information during processing, then moderate positive correlations would also be expected between probe accuracy and scores on the working memory tasks. Finally, if one of the causes of adult age differences in the matrix test is a reduced ability to preserve earlier presented information during subsequent processing, then older adults would be expected to be less accurate than young adults in these probe decisions.

Method

Subjects

Descriptive characteristics of the 30 young adults and 30 older adults who participated in this study are summarized in Table 4. The samples were recruited in the same manner described in Study 2, but none of the individuals had participated in a previous study concerned with either working memory or matrix reasoning.

Procedure

All subjects performed the tasks in the following order: Digit–Symbol; Digit–Digit; Reading Span; Computation Span; and Simultaneous Matrix. All tasks were identical to those in Study 2 except that the simultaneous matrix task also contained probes of the contents of previously displayed cells.

Prior to performing the matrix task on the computer the subjects studied a set of five practice problems (Numbers D1, D2, D3, C4 and E3 from the Raven's Standard Progressive Matrices) displayed in a booklet. Below each problem was a description of the attributes and relation relevant to the solution of the problem, and an explanation of why the designated answer was correct. The Simultaneous Matrix task was then described, with special emphasis on the probe trials.

Two blocks of 30 trials each were presented in the Simultaneous Matrix task. Within each block, 20 trials were standard trials in that the matrix was followed by the set of eight answer alternatives, and 10 trials were cell probes in which the matrix was followed by a probe concerning the contents of a particular cell. Within each stimulus set, each of the eight matrix cells was probed either once or twice, and the probes occurred on four trials with one relevant relation, four trials with two relevant relations, and two trials with three relevant relations.

The probes consisted of displays of a geometric pattern in one of the cells of the matrix accompanied by the words 'DIFFERENT' and 'SAME' on the lower left and lower right of the display, respectively. One half of the probes consisted of the same pattern presented earlier in that cell, and one half consisted of a different pattern (either the pattern from one of the other cells in the matrix, or a slightly altered version of the original pattern). Decisions to the probe stimuli were communicated by pressing the bottom left key ('Z') or DIFFERENT for the bottom right key ('/') for SAME.

Results and discussion

Descriptive statistics on several of the performance measures are summarized in Tables 4 and 5. As in Study 2, all the age differences were significant ($p < .01$) in the perceptual speed, working memory and matrix reasoning measures.

Results from the hierarchical regression analyses, summarized in the fifth row of Table 2, resembled those from Studies 1 and 2. Analyses restricted to subjects (22 young adults and 7 older adults) with accuracy of at least 75 per cent on problems with a single relation were also similar. In both analyses the variance in reasoning performance associated with age was substantially attenuated after statistical control of the composite measures of either working memory or perceptual speed.

Accuracy in the probe recognition trials averaged 70.3 per cent (SD = 11.6) for young adults and 57.3 per cent (SD = 12.5) for older adults ($t(58) = 4.18$, $p < .01$). An age × cell position analysis of variance was also conducted to determine whether the age differences varied as a function of the particular matrix cell whose contents were being probed. Both main effects in this analysis were significant (i.e. Fs > 6.4), as well as the age × cell position interaction ($F(7,406) = 4.28$, MSE = 849.47). Inspection of the cell means revealed that the interaction was attributable to both young and old adults performing near chance for one cell position (middle cell in the top row), but with young subjects performing more accurately than older subjects for all other cell positions. Examination of the stimulus patterns revealed that the DIFFERENT trials in this cell were created by rather subtle changes compared to those in the other cells, and thus an atypically difficult SAME/DIFFERENT discrimination may have been responsible for the low decision accuracy in top middle cell.

The correlation matrix containing the correlations among the major variables in this study is presented in Table 7. Notice that, as in the previous studies, the correlations among age, working memory and matrix reasoning performance are all moderately high. It should also be emphasized that the probe accuracy measure is positively correlated with both working memory and matrix reasoning, but negatively correlated with subject age.

Means of the percentage correct measure in each age group for problems with one, two and three relations are displayed in Fig. 8. The interaction of age × problem type in an analysis of variance was not significant (i.e. $F < 1$). Correlations between working memory and mean accuracy for problems with one, two and three relations were .60, .55 and .45, respectively. In both respects these results replicate those of Study 2.

The primary new results from this study concern the probe recognition measure. As expected, this measure was positively correlated with working memory and matrix reasoning performance, and negatively correlated with subject age. These results are consistent with the hypothesis that working memory mediates age differences in matrix reasoning tasks because of its role in preserving earlier information during subsequent processing.

Table 7. Correlation matrix for variables in Study 3 ($N = 60$)

Variable	1	2	3	4	5
1 Age	×	−.61*	−.50*	−.63*	.80*
2 SimAcc		(.86)	.67*	.61*	−.51*
3 ProbeAcc			(.47)	.50*	−.49*
4 WMem				(.71)	−.46*
5 Speed					(.92)
Mean	40.3	64.6	63.8	0.00	0.00
SD	21.0	17.2	13.6	0.88	0.96

* $p < .01$.

Note. Reliabilities in the diagonals were estimated by using the Spearman–Brown formula to boost the correlation between the two component measures (for Perceptual Speed and Working Memory) or by using the formula described by Kenney (1979, p. 132): reliability = $N(\text{avg.}\,r)/[1+(N-1)(\text{avg.}\,r)]$, with correlations between measures with one, two and three relations for the SimAcc and ProbeAcc variables.

Key. SimAcc = Accuracy in simultaneous condition of computer matrix task; ProbeAcc = Accuracy in cell recognition probes; WMem = Working memory composite created by averaging z scores from Reading Span and Computation Span tasks; Speed = Speed composite created by averaging z scores from Digit–Symbol and Digit–Digit tasks.

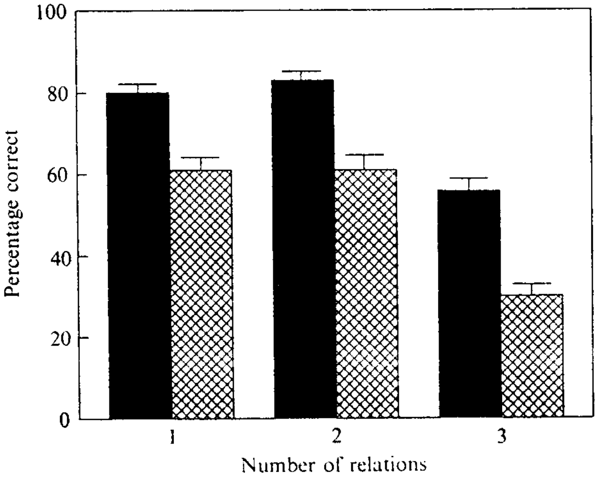


Figure 8. Mean accuracy for young and old adults as a function of the number of relations in the problem, Study 3. Lines above each bar correspond to one standard error. ■, young; ⊠, old.

STUDY 4

Although the results of Study 3 seem to implicate a mediational influence of working memory, it was somewhat surprising that accuracy of the probe recognition decisions was rather low in both age groups (i.e. 70.3 and 57.3 per cent for young and old adults, respectively). One interpretation of the relatively low levels of probe accuracy in the previous study is that information may be discarded when it is no longer necessary for processing. That is, because the probes in Study 3 were not presented until the subject indicated that he or she was ready to view the answer alternatives and make a decision, it is possible that the cell information had already been purged from memory by the time the probe was presented. This possibility can be investigated by interrupting the processing by the presentation of a probe about previously viewed information. In order to ensure that subjects were still engaged in processing at the time the probe was presented, the sequential version of the matrix task was used in this study.

The availability of measures of recognition probe accuracy and of the number of repetitive cell examinations in the inspection of the stimulus matrix also provides an opportunity to determine the relation between these two variables. If the redundant cell inspections occur because the previously viewed information is no longer available, then one might expect a negative correlation between the two measures because better preservation of information, as reflected by high probe accuracy scores, should be associated with a smaller number of repeated cell examinations.

Because accuracy in the sequential version of the matrix reasoning task in Study 2 was fairly low even among young adults (i.e. 70.7 per cent), an easier set of stimulus items was used in the present experiment. These were the problems used in an earlier study by Salthouse & Skovronek (1992), in which young adults averaged 90.7 per cent in the simultaneous conditions and 88.0 per cent in the sequential condition. In order to verify that performance on these new items was related to performance on the paper-and-pencil Raven's test, all subjects also performed the Raven's Advanced Progressive Matrices Test. However, because of a desire to keep the average length of the experimental session less than 1.5 hours, neither of the working memory tasks nor the Digit–Digit task were administered to subjects in this study.

Method

Subjects

Descriptive characteristics of the 30 young adults and 30 older adults who participated in this study are summarized in Table 4. None of the subjects had participated in any of the previous studies in this project, but each was recruited according to the procedures described in Study 2.

Procedure

All subjects began the session by performing the Raven's Progressive Matrices Test administered in the same manner described in Studies 1 and 2. They then performed both the standard WAIS-R Digit–Symbol Substitution Test, and the computer-administered digit–symbol version described earlier, followed by the sequential version of the computer matrix test.

Each subject performed three blocks of 18 trials each in the Sequential Matrix task, preceded by a block of six practice trials. The matrix problems in this study resembled those used in Studies 2 and 3, but were not constructed to differ systematically with respect to the number of relevant relations. Six trials in each experimental block were interrupted by the presentation of cell recognition probes. The probes were similar to those of Study 3, but were not identical because different sets of problems were used in this study. The particular matrix cell probed in the cell recognition trials varied randomly, but always occurred after two other cells had been examined since the previous viewing of the target cell. No recognition probe was presented if the subject did not examine the critical cell, or if he or she pressed the ENTER key to view the answer alternatives before examining two additional cells.

Results and discussion

Descriptive statistics summarized in Tables 4 and 5 indicate that, as in the previous studies, the young adults answered more Raven's items correctly, and were faster and more accurate in the computer matrix task then older adults. It can also be seen in Table 5 that although the young adults apparently benefited from the change to a different set of stimulus items (i.e. accuracy in the sequential condition was 70.7 per cent in Study 2 and 89.6 per cent in this study), the older adults still averaged less than 55 per cent correct decisions.

Subjects in this study also performed the paper-and-pencil WAIS-R Digit-Symbol Substitution Test. Scores on this test averaged 72.0 (SD = 9.8) for young adults, and 51.8 (SD = 12.1) for older adults ($t(58) = 7.12$, $p < .01$). These scores were converted to units of s per item by dividing them into 90 (the number of s allowed to perform the test), and then the z score for this measure was averaged with the z score from the computer-administered Digit–Symbol measure to form a composite perceptual speed variable.

Hierarchical regression analyses revealed that statistical control of perceptual speed reduced, but did not eliminate, the age differences in matrix reasoning performance. That is, the R^2 associated with age in the prediction of Raven's performance was .538 when considered alone, and .112 after control of perceptual speed. Corresponding values for the sequential matrix task were .630 and .094, respectively. These results are consistent with those of the previous studies, as summarized in Table 2.

Analyses of the cell examinations and of the probe accuracy measure revealed patterns similar to those reported in Studies 2 and 3. Specifically, young adults, compared to older adults, averaged fewer repetitive cell examinations (i.e. mean of 4.2, SD = 3.7 vs. mean of 7.6, SD = 7.3, $t[58] = 2.25$), and more accurate probe recognition decisions (i.e. mean of 84.8 per cent, SD = 12.8 vs. mean of 58.4 per cent, SD = 18.2, $t[58] = 6.49$). For the young adults, accuracy of the recognition probe decisions was greater in this study than in Study 3 (i.e. 84.8 vs. 70.3 per cent), but there was little difference in the accuracy of older adults across the two studies (i.e. 58.4 vs. 57.3 per cent).

The correlation matrix for the major variables in this study is presented in Table 8. The high correlation between the number of correct responses in the Raven's Progressive Matrices Test and accuracy in the sequential matrix test suggests that, as in Study 2, common processes contribute to performance in both types of tests. Also, like in Study 3, the moderately high correlations between the probe recognition accuracy measure and performance in both matrix reasoning tasks indicate that people who achieve high scores on the matrix reasoning tasks are also better than people with lower scores at preserving previously presented information.

It can be seen in Table 8 that the correlation between the averaged number of repetitive cell examinations and accuracy in the probe recognition trials was rather low (i.e. $-.21$), and not significantly different from 0. If the repetitive cell examinations represent an attempt to preserve information that is being lost, then one might have expected a large negative correlation. The failure to find more repetitive cell examinations among subjects with low probe accuracy scores casts

Table 8. Correlation matrix for variables in Study 4 ($N = 60$)

Variable	1	2	3	4	5	6
1 Age	×	−.73*	−.79*	−.70*	.32*	.79*
2 RPM		×	.81*	.60*	−.45*	−.67*
3 SeqAcc			(.94)	.74*	−.19	−.77*
4 ProbeAcc				×	−.21	−.64*
5 Repet.					×	.34*
6 Speed						(.85)
Mean	41.5	16.2	72.1	71.6	5.9	0.00
SD	22.7	6.4	25.5	20.5	6.0	0.92

* $p < .01$.

Note. Reliabilities in the diagonals were estimated by using the Spearman–Brown formula to boost the correlation between the two component measures (for Perceptual Speed) or by using the formula described by Kenney (1979, p. 132): reliability $= N(\text{avg. } r)/[1 + (N - 1)(\text{avg. } r)]$, with correlations between measures with one, two and three relations for the SeqAcc variable.

Key. RPM = Number correct in the Raven Progressive Matrices Test; SeqAcc = Accuracy in the sequential condition of the computer matrix task; ProbeAcc = Accuracy in cell recognition probes; Repet. = Average number of repetitive cell examinations in the sequential condition of the computer matrix task; Speed = Speed composite created by averaging *z* scores from paper-and-pencil and computer Digit–Symbol tasks.

doubt on this interpretation, although it is possible that the correlation was small because of low reliability of the measures. Unfortunately, no estimates of the reliability of the probe accuracy or repetitive cell examination measures were available in this study.

GENERAL DISCUSSION

The results of these studies replicated the findings of many earlier studies regarding substantial age-related declines on performance on the Raven's Progressive Matrices Test. The age correlation of −.57 in Study 1, and the discovery that the average older adult achieved scores equivalent to −3.6 (Study 2) and −2.6 (Study 4) young adult standard deviation units, are both consistent with the results summarized in the introduction.

The present results also extend the earlier results, however, by revealing that the age differences are evident in measures of accuracy, and do not simply reflect the number of items attempted. This was apparent in the analysis of the percentage of attempted items answered correctly in Study 1, and in the accuracy measures in each of the computer-administered matrix reasoning tasks. Older adults did take more time than young adults to inspect the matrices and to make their decisions, but they were also less accurate in those decisions. When expressed in young adult standard deviation units, the accuracy differences in the computer matrix tasks (cf. Table 5) averaged −2.27, −1.19, −1.60 and −4.88 units.

Another finding consistent with the results of other studies is that statistical control of a measure of working memory greatly attenuated the magnitude of the age-related effects on cognitive performance. The percentages by which the age effects were attenuated, as computed from the values in Table 2, were 83.5 per cent (Study 1) and 42.6 per cent (Study 2) for the Raven's Progressive Matrices Test, 68.9 per cent (Study 2) and 77.1 per cent (Study 3) for the simultaneous matrix task, and 71.1 per cent (Study 2) for the sequential matrix task. It therefore appears clear that working memory is an important factor in the age differences in these kinds of matrix reasoning tasks.

Three types of analyses were carried out to examine how working memory might mediate age differences in matrix reasoning. One set of analyses focused on between-item variation because of the assumption that at least some of the accuracy differences across items might be attributable to the demands placed on working memory. Contrary to expectation, however, the age effects were nearly constant across items varying considerably in mean accuracy. This is evident in Figs 3, 7 and 8 representing results from paper-and-pencil, simultaneous computer, and sequential computer administrations. A second set of analyses concentrated on error patterns. Examination of the frequency with which incorrect alternatives were selected (cf. Table 3) also failed to provide evidence of different patterns of erroneous responses. An implication of both sets of results is that the factors responsible for age differences are not the same as those contributing to the variation in item difficulty. As suggested in the discussion of Study 1, item variation in average accuracy may reflect factors such as the salience of the attributes and awareness of specific relational rules more than differential demands on working memory.

The third set of analyses designed to investigate how working memory might mediate adult age differences in matrix reasoning involved the probe recognition measures. If people with low levels of working memory do not function well in cognitive tasks because they are not very successful in preserving information while also carrying out processing, then they should be expected to be less accurate than people with higher working memory levels at recognizing probes of previously presented information. This hypothesis was supported in both Studies 3 and 4 as older adults were found to be significantly less accurate at recognizing the contents of previously examined matrix cells than young adults. Moreover, because the probe recognition measure was correlated with both the measures of working memory and the measures of cognitive performance, these results are compatible with the interpretation that one of the ways in which working memory exerts its mediating influence is by affecting the preservation of information during processing.

The finding of age differences in the accuracy of recognizing information presented earlier in the context of a cognitive task may have to be qualified because no age differences in a measure of probe recognition accuracy were reported by Salthouse & Skovronek (1992). A cube comparisons task was used in that project, and young and old adults were found to be equivalent in the accuracy of recognizing the contents of previously viewed cube faces. One possible explanation for this discrepancy is that age differences in the preservation of information may be pronounced only when the combined storage and processing requirements are demanding. That is, the matrix tasks in the current studies involved unique difficult-to-describe geometric patterns in each of eight cells, whereas the stimuli in the cube comparisons task consisted of only three to six familiar letters. Furthermore, Salthouse (1992*a*) recently found that the age differences in the accuracy of recognizing previously presented information in the context of an integrative reasoning task were greatly reduced by eliminating the data from subjects performing poorly in the simplest condition of the reasoning task. A similar result was evident in additional analyses of data from Studies 3 and 4 because employing a measure of accuracy with the simplest problems (i.e. one-relation problems in Study 3 and roughly equivalent problems in Study 4) as a covariate reduced the age differences in the probe recognition accuracy measure by approximately 50 per cent. The phenomenon of age differences in the recognition of information presented earlier in an on-going cognitive task may therefore vary as a function of the amount of information in the task, and of either the cognitive capabilities or the motivational involvement of the subjects.

An important remaining question concerns the factors responsible for the age differences in working memory. Previous research (Salthouse, 1991*a*; Salthouse & Babcock, 1991), including a report containing additional analyses incorporating the data from Studies 2 and 3 (Salthouse, 1992*a*), has revealed that the age differences in working memory are greatly attenuated by statistical control of measures of perceptual speed. The results summarized in Table 2 also indicate that the perceptual speed and working memory variables were nearly equivalent in the degree to which they influenced the age-related effects on matrix reasoning performance. That is, the values of R^2 for age after statistical control of working memory were very similar to those obtained after statistical control of perceptual speed, and in only a few cases was the attenuation of the age influence appreciably greater after control of both perceptual speed and working memory. These findings all suggest that an important determinant of the age differences in cognitive functioning, including measures of working memory, is the speed with which elementary cognitive operations can be executed. The mechanisms by which the rate of performing cognitive operations affects working memory and other cognitive tasks have not yet been identified, but the results of these and other studies indicate that the reductionistic analysis of age differences in cognition can, and should, be extended at least to a focus on speed of information processing as an explanatory construct.

Acknowledgements

This research was supported by National Institute on Aging Grant AG06826. I would like to thank R. Babcock, J. Jackson and J. Shaw for assistance in collection and analysis of data.

References

Carpenter, P. A., Just, M. A. & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, **96**, 404–431.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, **101**, 171–191.

Heron, A. & Chown, S. M. (1967). *Age and Function.* Boston: Little-Brown.

Jacobs, P. I. & Vandeventer, M. (1972). Evaluating the teaching of intelligence. *Educational and Psychological Measurement*, **32**, 235–248.

Jensen, A. R. (1982). The chronometry of intelligence. In R. J. Sternberg (Ed.), *Advances in the Psychology of Human Intelligence*, vol. I, pp. 255–310. Hillsdale, NJ: Erlbaum.

Jensen, A. R. (1983). Critical flicker frequency and intelligence. *Intelligence*, **7**, 217–225.

Jensen, A. R. (1987). Process differences and individual differences in some cognitive tasks. *Intelligence*, **11**, 107–136.

Kenney, D. A. (1979). *Correlation and Causation.* New York: Wiley.

Larson, G. E., Merritt, C. R. & Williams, S. E. (1988). Information processing and intelligence: Some implications of task complexity. *Intelligence*, **12**, 131–147.

Larson, G. E. & Saccuzzo, D. P. (1989). Cognitive correlates of general intelligence: Toward a process theory of g. *Intelligence*, **13**, 5–31.

Palmer, J., MacLeod, C. M., Hunt, E. & Davidson, J. E. (1985). Information processing correlates of reading. *Journal of Memory and Language*, **24**, 59–88.

Paul, S. M. (1985). The Advanced Raven's Progressive Matrices: Normative data for an American university population and an examination of the relationship with Spearman's g. *Journal of Experimental Education*, **54**, 95–100.

Raven, J. (1962). *Advanced Progressive Matrices, Set II.* London: H. K. Lewis.

Raven, J. C. & Court, J. H. (1989). *Manual for Raven's Progressive Matrices and Vocabulary Scales: Research Supplement no. 4.* London: H. K. Lewis.

Raven, J. C., Court, J. H. & Raven, J. (1985). *Manual for Raven's Progressive Matrices and Vocabulary Scales: General Overview.* London: H. K. Lewis.

Salthouse, T. A. (1990). Working memory as a processing resource in cognitive ageing. *Developmental Review*, **10**, 101–124.

Salthouse, T. A. (1991*a*). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science*, **2**, 179–183.

Salthouse, T. A. (1991*b*). *Theoretical Perspectives on Cognitive Aging.* Hillsdale, NJ: Erlbaum.

Salthouse, T. A. (1992*a*). Influence of processing speed on adult age differences in working memory. *Acta Psychologica*, **79**, 155–170.

Salthouse, T. A. (1992*b*). Working-memory mediation of adult age differences in integrative reasoning. *Memory & Cognition*, **20**, 413–423.

Salthouse, T. A. (1992*c*). Reasoning and spatial abilities. In F. I. M. Craik & T. A. Salthouse (Eds), *Handbook of Aging and Cognition.* Hillsdale, NJ: Erlbaum.

Salthouse, T. A. & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, **27**, 763–776.

Salthouse, T. A., Mitchell, D. R. D., Skovronek, E. & Babcock, R. L. (1989). Effects of adult age and working memory on reasoning and spatial abilities. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **15**, 507–516.

Salthouse, T. A. & Skovronek, E. (1992). Within-context assessment of age differences in working memory. *Journal of Gerontology: Psychological Sciences*, **47**, P110–P120.

Shipley, W. C. (1986). *Shipley Institute of Living Scale.* Los Angeles: Western Psychological Services.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale—Revised.* New York: Psychological Corporation.

Welford, A. T. (1958). *Aging and Human Skill.* London: Oxford University Press.

Received 30 May 1991; revised version received 17 June 1992