Memory & Cognition
1992, 20 (4), 413-423

Working-memory mediation of adult age differences in integrative reasoning

TIMOTHY A. SALTHOUSE
Georgia Institute of Technology, Atlanta, Georgia

Three research methods were used to investigate the hypothesized mediational influence of working memory on age-related differences in integrative reasoning. Results from all three procedures were consistent with the hypothesis because (1) statistical control of an index of working memory attenuated the age differences in reasoning accuracy, (2) young adults were more accurate than older adults in a measure reflecting the preservation of information during processing, and (3) young adults performing the task with a concurrent memory load exhibited a qualitative pattern of performance similar to that of older adults performing the task without a concurrent memory load.

Although there has recently been considerable interest in working memory as a possible mediator of adult age differences in cognition (for reviews, see Light, 1991; Salthouse, 1990; Stine & Wingfield, 1990), the amount of evidence directly relevant to the mediational role of working memory is still quite limited. The research reported in this article was therefore designed to use three investigative procedures to examine the hypothesis that working memory functions as a mediator of the age-related differences in at least some cognitive tasks.

The existence of negative relations between age and measures of fluid or process aspects of cognition has been well documented on the basis of research dating from the early decades of this century (see Salthouse, 1991b, for a review). Despite a considerable amount of research, no completely satisfactory explanation of these age-cognition relations is yet available. However, in the last few years there has been growing interest in the role of working memory as a proximal (i.e., assessed at the time of testing) mediator of age differences in at least some measures of cognitive functioning. One of the reasons for the enthusiasm about explanations based on working memory is that the construct of working memory seems more amenable to operationalization than do other relatively general constructs such as attentional capacity. That is, if one conceptualizes working memory as involving the simultaneous storage and processing of information, then it presumably can be measured by tasks in which there are both storage and processing requirements. For example, the two tasks used to operationalize working memory in the current project, computation span and reading span, involved subjects' retention of digit or word information while concurrently carrying out processing concerned with solving arithmetic problems or answering sentence-comprehension questions.

The focal cognitive task in this project was integrative reasoning involving abstract verbal material. A typical problem is illustrated in the top panel of Figure 1. Problems can vary in the number of premises describing relations between two terms, but all eventually end with a question asking what will happen to one term if a specified change is introduced in another term. An advantage of this particular task is that problem difficulty (as reflected by decision accuracy) can be systematically altered by manipulation of the number of premises presented prior to the question, presumably because demands on working memory are increased by the additional requirements for the storage and coordination of information associated with more premises.

The three investigative procedures used in this project were all designed to examine the hypothesis that age-related differences in the integrative reasoning task are at least partially attributable to age-associated reductions in some aspect of working memory. The statistical control procedure tests the prediction that the magnitude of the age differences should be greatly reduced if people of different ages are statistically equated with respect to their working-memory ability. Because working memory involves the preservation (storage) of information during processing, the experimental analysis procedure examines the implication that older adults should be less accurate than young adults at maintaining relevant information while engaged in the performance of the task. And finally, if a low level of working memory can be considered analogous to a concurrent memory load, then the pattern of performance exhibited by older adults should be qualitatively similar to that found in young adults who perform the task while also remembering other information. This expectation is examined in the simulation procedure.

This research was supported by NIA Grant R37 AG06826. I would like to thank J. Jackson and A. Kersten for assistance in testing subjects, and the members of the Cognitive Aging Group at Georgia Tech (R. Babcock, J. Earles, C. Hertzog, A. Mattina, W. Rogers, B. Rypma, and A. Smith) for helpful comments on an earlier draft of this manuscript. Correspondence should be addressed to T. A. Salthouse, School of Psychology, Georgia Institute of Technology, Atlanta, GA 30332-0170.

(All Relevant)

R and S do the OPPOSITE
Q and R do the SAME
If Q INCREASES, what will happen to S?

(1 Relevant, 1st position)

G and H do the SAME
F and G do the OPPOSITE
If G DECREASES, what will happen to H?

(Recognition Probe, 1st position)

D and E do the SAME
E and F do the OPPOSITE
*** D and E do the SAME ***

Figure 1. Illustration of sample problems in the integrative-reasoning task.

Statistical control procedures have been used in two earlier research projects with this same integrative-reasoning task. Salthouse, Mitchell, Skovronek, and Babcock (1989) administered the reasoning task and the computation-span working-memory task to 120 adults between 20 and 80 years of age. Multiple-regression equations predicting reasoning accuracy revealed that age was associated with an R^2 of .278 when considered alone, but the R^2 value for age was reduced to only .119 after partialing the variance associated with the working-memory measure. Two additional studies, each involving over 220 adults between 20 and 80 years of age who were administered the integrative-reasoning task and two working-memory tasks, were reported by Salthouse (1991a). The working-memory tasks in these studies were group-administered versions of the computation-span and listening-span tasks (see Salthouse & Babcock, 1991, for further description). Results similar to those of Salthouse et al. (1989) were obtained in both studies. To illustrate, values of R^2 associated with age in the prediction of reasoning accuracy were .121 and .104 before statistical control of a composite measure of working memory, but were only .015 and .036 after such control. The results of these earlier studies, therefore, suggest that a fairly large proportion of the age-related differences in the integrative-reasoning task may be mediated by reductions in working memory. In terms of percentages, the variance associated with age was reduced by 57.2%, 87.6%, and 65.4%, respectively, across the three studies.

Although the statistical control procedure can be very informative, it does have several limitations. For example, the fact that only a single measure of working memory was used in the Salthouse et al. (1989) study raises the possibility that the assessment of working memory was relatively narrow and closely linked to particular types of information and specific kinds of processing operations. The Salthouse (1991a) studies included two distinct working-memory measures, but the group-administration format may have created a spurious association between the working-memory and reasoning measures because of factors unrelated to working memory (e.g., time restrictions necessitated by group testing). An attempt was made to minimize the possibility of an artifactual relation between the working-memory and reasoning measures in the present study by administering the tasks by means of computers. This allowed research participants to spend as much time as desired in each phase of the working-memory tasks and to respond at their own pace in the reasoning task.

If working memory is important for cognitive tasks because it allows information to be preserved while other information is being processed, then one should expect to find better preservation of information during the performance of a cognitive task among people whose working-memory systems are highly effective than among people whose working-memory systems are less effective. This prediction can be tested by using versions of the experimental analysis procedure to evaluate the availability of information presented earlier in an ongoing cognitive task. Information availability can be examined with a variety of different methods, but the critical point is that the evidence for age differences in working memory is derived while the subject is actively engaged in the performance of a cognitive task and not from a separate task deliberately designed to assess working memory. In the terminology of Salthouse (1990), measures obtained with the experimental analysis procedure are within-context measures, whereas those typically used in the statistical control procedure are out-of-context measures.

Previous research (e.g., Salthouse, Legg, Palmon, & Mitchell, 1990; Salthouse et al., 1989) with the current integrative-reasoning task examined information availability in terms of the contrast between two types of trials. Trials in which all premises are relevant to the decision (illustrated in the top panel of Figure 1) require both memory and integration of information. That is, because the question concerns terms originally mentioned in different premises, information from several premises must be integrated and coordinated to determine the answer. In contrast, no information integration is required in trials in which only one premise is relevant because the question concerns terms mentioned in a single premise. An example is illustrated in the middle panel of Figure 1, in which the question refers to two terms originally related in the first premise. Because trials in which all premises are relevant require both memory and integration whereas those with a single relevant premise do not require across-premise integration, it has been assumed that comparisons of accuracy in the two types of trials would be informative about the relative importance of memory and integration processes in the performance of this particular task.

In both earlier studies (Salthouse et al., 1989, 1990), decision accuracy decreased as more premises were presented, and the magnitude of accuracy decline for trials with one relevant premise was nearly the same as that for trials with all relevant premises. Furthermore, the same general pattern was evident in adults of all ages. These results were interpreted as suggesting that a critical factor in the variation in accuracy as a function of number of premises and as a function of adult age may be the preservation of information, and not its integration or coordination. That is, the principal determinant of the variations in reasoning accuracy associated with additional premises or with increased age appears to be the ability to maintain previously presented information rather than the ability to integrate information that is available in memory.

As with almost any experimental technique, however, objections can be raised with respect to how results of the manipulations are best interpreted. One reservation about the contrast between one-relevant and all-relevant trials is that even performance on one-relevant trials may not be a pure measure of information availability, but might also reflect the ability to transform the format of the information from the terms "opposite" and "same" used in the premises to the terms "increase" and "decrease" used in the question. Rather than representing the ability to maintain information without integration, therefore, accuracy of decisions in one-relevant trials could reflect the ability to maintain information *and* to convert it from one representational format to another.

This concern can be addressed by attempting to obtain a more direct assessment of information availability in the form of recognition probes for previously presented premise information. An example of a probe-recognition trial used in the current experiment is illustrated in the bottom panel of Figure 1. Notice that the probe is presented in the same format as the premises, and the subject merely has to decide whether it is the same as, or different from, one of the premises presented earlier in that trial. Supplementing the contrast of performance in one-relevant and all-relevant trials with these recognition probes should therefore allow a more definitive evaluation of the hypothesis that older adults perform poorly on integrative-reasoning tasks because they are less likely to preserve relevant information than are young adults. Because recognition probes provide a more direct assessment of the status of information in memory, a finding of age differences in measures of probe-recognition accuracy would provide additional support for the inference that a failure to maintain critical information during processing is an important determinant of age differences in integrative-reasoning performance.

The principal assumptions underlying the simulation procedure are that the effects of reduced working-memory functioning can be mimicked by various experimental manipulations, and hence that the pattern of performance exhibited by young subjects performing the reasoning task with one of those manipulations should be qualitatively similar to the pattern exhibited by older subjects performing the task under normal conditions. Baddeley and Hitch (1974; Hitch & Baddeley, 1976) have proposed that requiring subjects to maintain a concurrent memory load reduces effective working-memory functioning because some aspects of working memory have to be devoted to preserving the task-irrelevant memory-load information. If this assumption is correct, and if a major factor contributing to the age differences in reasoning performance is a diminished working memory among older adults relative to young adults, then the patterns of performance for older adults should be expected to be very similar to those for young adults performing the task with a concurrent memory load. Both groups should have lower levels of performance than young adults without a concurrent memory load, and the pattern of performance differences between young adults and old adults should resemble that between young adults with and without the concurrent memory load.

The major limitation of the simulation approach is that direct quantitative comparisons are generally not meaningful when there is no basis for specifying the correspondence between a particular number of years of age difference and a given amount of concurrent memory load. Qualitative comparisons evaluated with respect to a common reference group are also complicated because impairments in performance can be produced in a variety of different ways. Confidence that the simulation is accurate thus tends to increase in proportion to the number and pattern of correspondences that can be established. Fortunately, the reasoning task investigated in this project allows several possible contrasts between the performance of older adults and that of young adults with a concurrent memory load by comparing each against the performance of young adults with respect to (1) the effect of additional premises on decision accuracy, (2) the pattern of accuracy of one-relevant trials relative to all-relevant trials, (3) the pattern of accuracy of one-relevant trials as a function of the serial position of the relevant premise, and (4) the pattern of accuracy of recognition-probe trials as a function of the serial position of the relevant premise.

To summarize, if an important factor contributing to adult age differences in integrative reasoning is an age-related reduction in working memory, then one would expect (1) substantial attenuation of the age differences in reasoning after statistical control of measures of working memory, (2) significantly lower accuracy for older adults compared with young adults in measures of information availability obtained while subjects are engaged in the performance of the task, and (3) a qualitatively similar pattern of performance for older adults and for young adults performing the task with a concurrent memory load. The study described below was designed to test each of these predictions.

In addition to the primary focus on working memory, there was also a secondary interest in the role of perceptual-comparison speed as a potential mediator of the age differences in both reasoning and working memory. Previous studies have revealed that statistical control of an index of perceptual speed greatly attenuated the age differences in measures of reasoning and working memory (Salthouse, 1991a; Salthouse & Babcock, 1991). However, the same paper-and-pencil measures of perceptual-comparison speed were used in all of the earlier studies, and consequently the generalizability of this result is not yet known. Two computer-administered measures of perceptual-comparison speed based on the Digit Symbol Substitution Test (Salthouse, 1992) were therefore employed to examine the influence of perceptual speed on the relations between age and performance on both the reasoning and working-memory tasks.

METHOD

Subjects

Characteristics of the three samples of adults who participated in this project are summarized in Table 1. Older adults were recruited from newspaper advertisements and referrals from other participants and were paid a nominal fee for their participation. Both samples of young adults were college students who received extra credit in psychology courses for their participation. It is apparent in Table 1 that the three groups were similar with respect to average years of education and self-rated health, and thus any age-related differences in the performance measures are unlikely to be confounded with differences in at least these characteristics.

Procedure

All participants performed four tasks, in the same order, before the reasoning task. The first task, digit-symbol, was a computer-administered version of the WAIS-R Digit Symbol Substitution Test (Wechsler, 1981). Displays in this task consisted of a code table associating digits and symbols and a single test stimulus consisting of a digit-symbol pair. The code table remained constant across trials, but the identity of the digits and symbols in the test stimulus varied from trial to trial. The task was to press the / (slash) key on the keyboard as rapidly as possible when the digit and symbol in the test stimulus matched according to the code table and to press the Z key on the keyboard when they did not match. The digit-digit task was very similar but had no symbols. The code table was therefore uninformative in this task because the decisions were to be made on the basis of the physical identity of the pair of digits. In both tasks, a practice set of 18 trials was followed by an experimental set of 90 trials. Because accuracy was greater than 95% in both tasks for all three groups, the median time per response in the experimental trials served as the dependent measure.

The two working-memory tasks required subjects to remember information while also carrying out specified processing. The to-be-remembered information in the reading-span task was the last word in the sentence, and the processing consisted of selecting which of three alternatives was the correct answer to a simple question about the sentence. The to-be-remembered information in the computation span task consisted of the second digit in an arithmetic problem, and the processing consisted of selecting which of three alternatives was the correct answer to the problem. Materials in both tasks were identical to those described in Salthouse and Babcock (1991). Each successive item (sentence or arithmetic problem) could be viewed as long as desired, and the subjects were allowed as much time as needed to type their recall responses. A trial consisted of the presentation of the designated number of sentences or arithmetic problems followed by an attempt to recall the last items from all of the presented sentences or arithmetic problems in the order in which they were presented. Spans were determined by the longest sequence in which the target items were recalled correctly and the processing was performed without mistakes on at least two of the three trials. A practice phase with trials containing sequences of up to three items preceded the experimental phase in each task.

The reasoning task was first described, then administered in a practice block of seven trials. Two experimental blocks of 63 trials each were then presented. All participants received the same trials in the same sequence. Forty-five of the trials were normal reasoning trials, and 18 were recognition-probe trials. Across the two blocks there were an equal number of *increase* (or *same*) and *decrease* (or *different*) trials for each of the five combinations of number and type of premises (i.e., 1-1, 2-1, 2-2, 3-1, and 3-3, where

Table 1
Means and Standard Deviations for Variables in the Three Samples

	Young		Old		Young—Memory Load	
N	30		30		30	
Proportion females	.50		.53		.30	
	M	*SD*	*M*	*SD*	*M*	*SD*
Age	20.1	1.0	68.2	5.7	20.3	1.5
Health	1.5	0.7	1.7	0.9	1.6	0.6
Education	14.2	1.1	14.5	2.2	14.1	1.3
DS time	1.13	0.14	1.91	0.36	1.19	0.19
DD time	0.53	0.05	0.77	0.27	0.53	0.04
CSpan	4.4	2.1	2.2	1.7	4.7	1.8
RSpan	2.9	1.5	1.9	1.6	3.3	1.8
Reas. acc.	79.1	11.1	59.9	9.2	63.6	11.8
Probe acc.	77.3	15.2	66.4	12.2	68.6	15.6

Note—Health is self-rating on a scale where 1 = *excellent* and 5 = *poor*. Education is self-reported years of formal education completed. DS time is time in seconds in digit-symbol task. DD time is time in seconds in digit-digit task. CSpan is the span in the computation-span task. RSpan is the span in the reading-span task. Reas. acc. is percentage correct in the normal-reasoning trials. Probe acc. is percentage correct in the probe-recognition trials.

the first digit refers to the total number of premises and the second digit to the number of relevant premises). All types of trials, including the recognition-probe trials, were randomly intermixed within the experimental block, but all participants received the same random sequence.

Premises were presented sequentially for 4 sec each. The question remained visible until the subject's response (Z for *increase*, / [slash] for *decrease*), but the instructions emphasized that the responses should be made as rapidly as was consistent with maximum accuracy.

Recognition-probe trials were identical to the reasoning trials except that the recognition probe replaced the question requiring an *increase* or *decrease* decision. One half of the probe trials had the same premise as one of those presented earlier in the trial, and one half had a premise that differed from an earlier premise by a reversal of the original relation (e.g., a premise with a *same* relation was changed to one with an *opposite* relation). Decisions that the premise was identical to one presented earlier were communicated by pressing the / (slash) key, and decisions that the premise was different were communicated by pressing the Z key.

Trials for the subjects in the concurrent-memory-load condition were preceded by the presentation of a set of five randomly selected digits. The subjects viewed the digits for 4 sec, after which the first display of the reasoning task was presented. Immediately following the decision response to the reasoning question or recognition probe, the subject was asked to recall the five digits in the order in which they were presented by typing them on the computer keyboard. Performance in the reasoning tasks was only analyzed for trials in which the subjects were correct on the digit recall. This included 101 trials for the average subject, with a range across subjects of 48 to 125.

RESULTS

Statistical Control Analyses

Multiple-regression analyses were conducted to examine the statistical control predictions of attenuated age relations on the measure of reasoning accuracy after partialing the variance associated with working memory. Because measures of information availability, in the form of probe-recognition accuracy, and of perceptual speed, from the digit-symbol and digit–digit tasks, were obtained from each participant, these measures were also entered into the regression analyses. Composite measures of working memory and perceptual speed were created by averaging the z scores from the measures in the computation span and reading span tasks to create a working-memory composite and averaging the z scores from the measures in the digit-symbol and digit–digit tasks to create a perceptual-speed composite. (Correlations between the measures in the total sample after partialing out age were .38 between the two working-memory measures and .56 between the two perceptual-speed measures. The remaining correlations, again after partialing age, were −.09 between computation span and digit-symbol, −.07 between computation span and digit–digit, −.16 between reading span and digit-symbol, and −.04 between reading span and digit–digit.) Results of the regression analyses based on the complete samples of young and old adults, and of young adults with and without a concurrent memory load, are summarized in the top panel of Table 2.[1]

Entries in the first column of Table 2 refer to the cumulative R^2 after the variable in that row and the immediately preceding rows had been entered into the regression equation. Values in the second column indicate the increment in R^2 when the variable in that row is added to the regression equation. The values in the third column are F ratios evaluating the statistical significance of the initial R^2 or the increment in R^2 associated with the added variable. Finally, the entries in the fourth column indicate the percentage by which the group differences were attenuated by statistical control of the preceding variables. For example, the difference between the initial R^2 for age of .479 and the increment in R^2 for age after controlling the influence of working memory of .248 is .231, which corresponds to an attenuation of 48.2%.

Because some subjects had an average accuracy very close to chance, the analyses were repeated after omitting subjects with less than 80% correct responses on the reasoning trials with only one premise. This is an arbitrary criterion, but it does ensure that the level of understanding and degree of motivation was sufficiently high that the task could be performed with moderate accuracy in the simplest condition. These data, based on 26 young adults, 16 older adults, and 12 young adults with a concurrent memory load, are displayed in the bottom panel of Table 2.

Three points should be noted about the results of the analyses involving the groups of young adults without a concurrent memory load and old adults. The first is that the relative amount of attenuation (i.e., 48.2% in the complete sample and 32.4% in the restricted sample) of the age-associated variance after control of working memory is similar to that found in the previous studies using the statistical control procedure (e.g., Salthouse et al., 1989; Salthouse, 1991a).

The second point is that the reduction in age-related variance associated with control of the probe-accuracy measure is approximately the same as that associated with the working-memory measure (i.e., 49.3% vs. 48.2% in the complete sample) and that the additional reduction when both measures are controlled simultaneously is rather small (i.e., to 61.8%). This suggests that, at least with respect to the age-related differences on reasoning accuracy, the two variables are not independent but instead have a certain amount of shared or common variance.

Finally, it can be seen that statistical control of the perceptual-speed measure, either by itself or in combination with other variables, resulted in an attenuation of the age-related variance by 80% to 90%. The apparent implication is that a large proportion of the age-related differences in this reasoning task is mediated by age-related reductions in speed of processing, as indexed by the perceptual-speed measures.

Results from the contrasts of young adults with and without a concurrent memory load are presented in Table 2 primarily for purposes of comparison with the age contrasts. Notice that unlike the differences between

Table 2
Results of Hierarchical Regression Analyses on Average Reasoning Accuracy

	Young vs. Old				Young vs. Young—Memory Load			
	R^2	Increment in R	F	Percent Attenuated	R^2	Increment in R^2	F	Percent Attenuated
			All Subjects					
Group	.479		53.28*		.321		27.44*	
Working memory	.283		34.41*		.009		0.77	
Group	.531	.248	30.12*	48.2	.346	.337	29.41*	−5.0
Probe accuracy	.396		62.50*		.461		68.56*	
Group	.639	.243	38.27*	49.3	.617	.156	23.19*	51.4
Working memory	.283		44.74*		.009		1.31	
Probe accuracy	.463	.180	28.35*		.461	.452	67.01*	
Group	.646	.183	28.92*	61.8	.622	.161	23.87*	49.8
Perceptual speed	.466		58.55*		.100		9.08*	
Group	.547	.081	10.21*	83.1	.370	.270	24.38*	15.9
Perceptual speed	.466		78.22*		.100		15.03*	
Probe accuracy	.605	.139	23.36*		.480	.380	56.77*	
Group	.667	.062	10.43*	87.1	.626	.146	21.86*	54.5
Perceptual speed	.466		62.31*		.100		9.09*	
Working memory	.529	.063	8.43*		.101	.001	0.06	
Group	.582	.053	7.10	88.9	.382	.281	25.40*	12.5
Perceptual speed	.466		77.93*		.100		14.87*	
Working memory	.529	.063	10.54*		.101	.001	0.10	
Probe accuracy	.620	.091	15.28*		.480	.379	56.13*	
Group	.671	.051	8.64*	89.4	.629	.149	22.03*	53.6
			Subjects Satisfying Accuracy Criterion					
Group	.426		29.72*		.147		6.35	
Working memory	.184		13.62*		.122		6.03	
Group	.472	.288	21.26*	32.4	.270	.148	7.28	−0.1
Probe accuracy	.166		13.38*		.312		20.33*	
Group	.517	.351	28.45*	17.6	.448	.136	8.88*	7.5
Working memory	.184		14.94*		.122		8.25*	
Probe accuracy	.254	.070	5.67		.344	.222	14.94*	
Group	.531	.277	22.44*	35.0	.482	.138	9.31*	6.1
Perceptual speed	.451		34.52*		.117		5.33	
Group	.490	.039	3.03	90.8	.211	.094	4.30	36.1
Perceptual speed	.451		36.49*		.117		7.66*	
Probe Accuracy	.525	.074	4.36		.361	.244	16.02*	
Group	.564	.039	2.06	90.8	.466	.105	6.86	28.6
Perceptual speed	.451		39.32*		.117		5.89	
Working memory	.505	.054	6.45		.200	.083	4.19	
Group	.530	.025	3.43	94.1	.305	.105	5.28	
Perceptual speed	.451		39.46*		.117		7.85*	
Working memory	.505	.054	4.72		.200	.083	5.59	
Probe accuracy	.547	.042	3.71		.384	.184	12.33*	
Group	.577	.030	2.62	93.0	.494	.110	7.36	25.2

*$p < .01$.

young and old adults, the performance differences associated with the memory-load condition were not substantially attenuated by statistical control of the working-memory or perceptual-speed variables. This is to be expected because the two young adult groups were very similar in these out-of-context measures (cf. Table 1).

Analyses of Information Availability

Accuracy in the reasoning task as a function of the number of presented premises is displayed in Figure 2. The top panel contains the data from all subjects, and the bottom panel presents results from the subset of subjects in each group with accuracy of at least 80% on trials with only one premise. Notice that accuracy decreases with more premises and that the absolute level of performance is very similar for trials with one relevant premise and for trials in which all premises were relevant. The analysis of variance (ANOVA) results, summarized in Table 3, confirm that the main effects of group and number of premises were significant but that the relevance effect

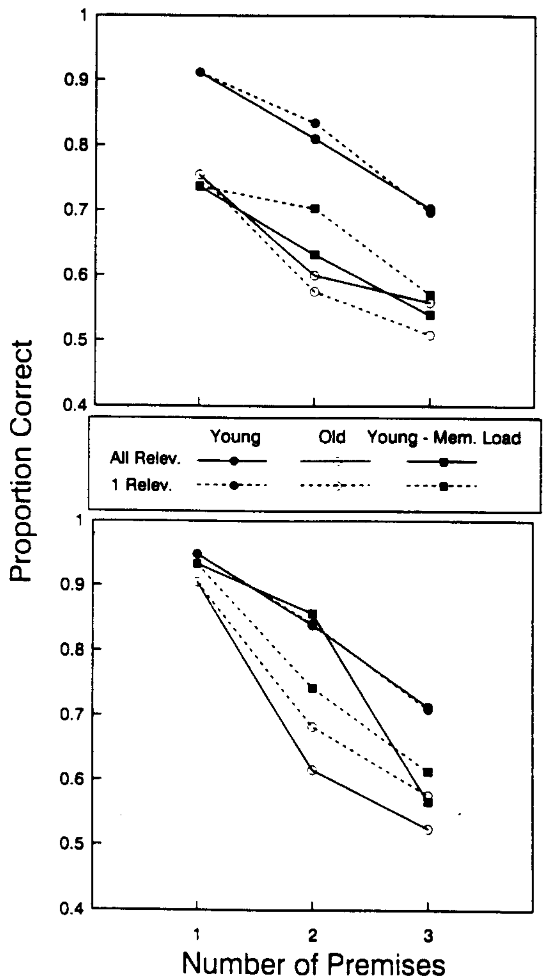

Figure 2. Accuracy in trials with all premises relevant to the decision or with only one premise relevant to the decision as a function of the number of presented premises. The top panel presents data from all subjects, and the bottom panel shows data from subjects with at least 80% accuracy on trials with a single premise.

(i.e., one relevant vs. all relevant) was not significant. The primary difference between the patterns in the top and bottom panels of Figure 2 is that in the bottom panel the differences between the young and old groups appear to increase as additional premises are presented. This pattern was verified in the ANOVA as the interaction of age group × number of presented premises was significant in the data from the restricted sample. With this single exception, the results from the analyses involving the contrast between young adults with and without the concurrent-memory-load requirement were similar, particularly with respect to the absence of interactions involving the group factor.

Accuracy was also analyzed on one-relevant trials as a function of the serial position of the relevant premise. These analyses were of interest because there were several ways in which average accuracy could have decreased with the presentation of additional premises. For example, accuracy might have remained invariant for either the first or the last premise in the trial but declined on other premises, or it might have declined nearly uniformly across all premise positions. The results, displayed in Figure 3, indicate that the second characterization is more accurate. That is, for all three groups, accuracy was lower with more premises regardless of the serial position of the relevant premise. Furthermore, young adults without a concurrent memory load were always more accurate than were subjects in the other two groups. Results of the ANOVAs with age (young, old) and premise type (i.e., 1-1, 2-1, 2-2, 3-1, 3-2, and 3-3, where the first and second digits refer to number of premises and position of relevant premise, respectively)[2] as factors are summarized in Table 4. Of particular interest is the absence of an age × premise type interaction in both the analysis based on all of the data and in the analysis based on the data from the restricted sample. Virtually identical results were evident in the analyses contrasting the two young adult groups distinguished by the presence or absence of a concurrent memory load.

Figure 3 also contains probe-recognition accuracy plotted as a function of serial position of the probed premise. The general pattern closely resembles that of the one-relevant reasoning trials except when the trials contained two premises, where the older adults and the young adults with a concurrent memory load were somewhat more accurate in the recognition trials than in the reasoning trials. Results of the age group × premise type ANOVAs similar to those conducted on the accuracy of one-relevant reasoning trials are summarized in Table 5. The major point to be emphasized from these data is that the age difference in recognition accuracy was eliminated by restricting the sample to subjects with accuracy of at least 80% in the reasoning trials with a single premise. To verify that the elimination of the age differences was not simply attributable to the reduced power associated with a smaller sample size, an analysis of covariance with accuracy in one-premise reasoning trials as the covariate was also conducted. The age group main effect was also not significant in this analysis [$F(1,57) = 0.85$, $MS_e = 131.06$], thus confirming the results from the restricted sample analysis.

Qualitative Similarity of Age and Concurrent-Memory-Load Contrasts

Inspection of Figures 2 and 3 and Tables 3, 4, and 5 reveals that the performance of all three groups was qualitatively very similar. This is evident in the declines in accuracy with additional premises, in the nearly equivalent levels of accuracy for one-relevant and all-relevant trials, and in the serial-position functions for one-relevant reasoning trials and for probe-recognition trials. Statistical support for this similarity is evident in the general absence of interactions involving the group variable in the ANOVA results summarized in Tables 3, 4, and 5.

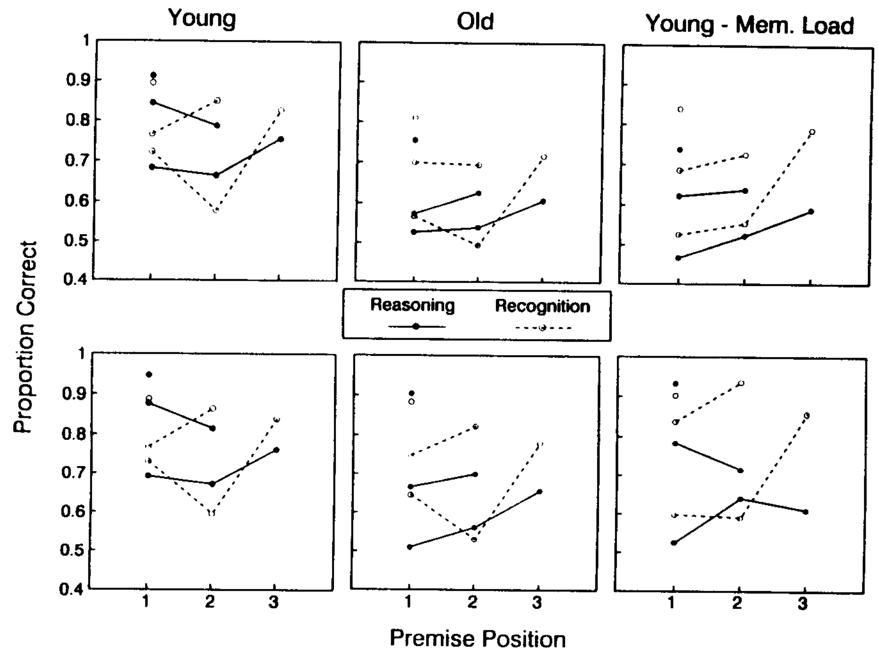

Figure 3. Accuracy in reasoning trials with one relevant premise and in recognition-probe trials as a function of the serial position of the relevant or probed premise. The top panel presents data from all subjects, and the bottom panel shows data from subjects with at least 80% accuracy on trials with a single premise.

DISCUSSION

The results of this study suggest that there are at least two factors involved in the age differences in integrative-reasoning performance. One factor is fairly broad, in that it affects performance in all conditions of the task and may reflect lack of understanding, low motivation, or some type of sensory or motor deficit. It is assumed that the influence of this factor can be minimized by eliminating subjects with low levels of performance in the simplest trials involving only one premise. The magnitude of the age differences when the sample was restricted in this manner was somewhat smaller than that in the complete sample, but there were only two cases where the pattern of results appeared different. One difference occurred in the analyses summarized in Table 3, where the interaction

Table 3
Results of Analyses of Variance on Reasoning Accuracy

	Young vs. Old			Young vs. Young—Memory Load		
	df	*F*	MS_e	*df*	*F*	MS_e
		All Subjects				
Group	1,58	47.50*	658.24	1,58	27.68*	815.04
Number of premises	2,116	69.87*	205.45	2,116	64.14*	184.98
Group × number	"	2.63	"	"	0.37	"
Relevance (one vs. all)	1,58	0.60	135.19	1,58	2.93	122.94
Group × relevance	"	1.60	"	"	1.43	"
Number × relevance	2,116	0.72	109.62	2,116	1.43	125.88
Group × number × relevance	"	0.50	"	"	0.36	"
		Subjects Satisfying Accuracy Criterion				
Group	1,40	28.29*	363.41	1,37	6.16	422.86
Number of premises	2,80	140.00*	125.33	2,74	107.33*	133.41
Group × number	"	9.69*	"	"	3.51	"
Relevance (one vs. all)	1,40	1.74	133.07	1,37	0.70	114.90
Group × relevance	"	1.83	"	"	0.64	"
Number × relevance	2,80	0.65	97.40	2,74	2.84	104.59
Group × number × relevance	"	0.64	"	"	3.20	"

$^*p < .01$.

Table 4
Results of Analyses of Variance on Reasoning Accuracy with One-Relevant Premise

	Young vs. Old			Young vs. Young—Memory Load		
	df	*F*	MS_e	*df*	*F*	MS_e
			All Subjects			
Group	1,58	33.93*	765.62	1,58	29.81*	961.98
Premise type	5,290	13.85*	314.24	5,290	14.52*	357.24
Group × premise type	"	1.23	"	"	0.46	"
			Subjects Satisfying Accuracy Criterion			
Group	1,40	16.39*	590.68	1,37	8.86*	635.42
Premise type	5,200	19.20*	292.80	5,185	15.37*	334.63
Group × premise type	"	1.22	"	"	0.92	"

*$p < .01$.

Table 5
Results of Analyses of Variance on Probe-Recognition Accuracy

	Young vs. Old			Young vs. Young—Memory Load		
	df	*F*	MS_e	*df*	*F*	MS_e
			All Subjects			
Group	1,58	9.40*	1139.93	1,58	4.77	1422.07
Premise type	5,290	22.12*	339.29	5,290	22.16*	357.45
Group × premise type	"	0.66	"	"	1.71	"
			Subjects Satisfying Accuracy Criterion			
Group	1,40	1.15	985.06	1,37	0.02	1214.24
Premise type	5,200	18.62*	288.88	5,185	17.75*	310.53
Group × premise type	"	0.37	"	"	1.14	"

*$p < .01$.

of age × number of premises was significant in the restricted sample but not in the complete sample. This pattern, illustrated in the bottom panel of Figure 2, is consistent with earlier findings of larger age differences with a greater number of premises (Salthouse et al., 1989, 1990). Inclusion of data from subjects who performed at or near chance levels in the task apparently obscured this interaction in the analyses based on data from the complete sample.

The second difference in the results of the complete sample and the restricted sample was in the analysis of probe-recognition accuracy, summarized in Table 5. In this case, the age differences were no longer significant when subjects performing poorly in the simplest reasoning problems were eliminated from the analyses. This pattern, which is supported by the absence of significant age differences in the analysis of covariance, suggests that, at least when there is some assurance that everyone understands the task and is apparently motivated to perform, young and old adults are equivalent in the ability to preserve untransformed information during processing. A similar result in a cube-comparison task was reported by Salthouse and Skovronek (1992), where it was concluded that age differences in working memory are pronounced only when the stimulus information has to be manipulated or transformed in some fashion. This interpretation may also apply in the present study because transformation (from *same* or *opposite* to *increase* or *decrease*) is required with the reasoning questions, but not with the recognition probes. The tendency for young adults to have nearly the same accuracy in the one-relevant reasoning trials and the probe-recognition trials (i.e., 79.4% vs. 77.9%) but for older adults to be less accurate in the reasoning trials (i.e., 66.7% vs. 73.6%) is also consistent with this interpretation.[3]

No single set of results is definitive, but the combined results from three quite different procedures appear fairly convincing in suggesting that working memory is a major factor contributing to adult age differences in integrative reasoning. That is, working memory seems to be implicated because of the substantial attenuation of the age differences after statistical control of an index of working memory, the presence of age differences in at least some measures of information availability obtained during the performance of the task, and the qualitatively similar pattern of performance differences between young adults and old adults and between young adults performing under normal and under concurrent-memory-load conditions. Each of these procedures has limitations, but because they are not the same limitations, confidence in one's inferences is enhanced when the results from each procedure converge on the same interpretation. It therefore seems reasonable to conclude that one of the causes of adult age differences in certain cognitive tasks is a limited ability to preserve information during processing, which can be viewed as a consequence of an impairment in working memory.

The fact that the premises in the reasoning task were presented sequentially, and for a duration of only 4 sec, raises the possibility that there was a greater influence of working-memory or perceptual-speed factors in this task than in more traditional reasoning tasks. This interpretation is plausible, but it should be noted that similar patterns of age differences, and attenuations of those differences after statistical control of an index of working memory, were observed in previous studies in which the premises either were presented sequentially under self-paced conditions (Salthouse et al., 1989) or were all presented simultaneously in a paper-and-pencil format (Salthouse, 1991a). It therefore seems unlikely that the results of the current study have limited generalizability because of the specific method used to present the stimulus materials.

If it is the case that age-related reductions in working memory play an important role in the age differences in this, and perhaps other, cognitive tasks, the question immediately arises as to what is responsible for age-related differences in working memory. Results from other research suggest that the speed of executing relatively simple operations probably contributes to the age differences in working memory. As an example, Salthouse and Babcock (1991) found that a large proportion of the age differences in two measures of working memory were reduced after statistical control of measures of perceptual speed, and these findings were replicated and extended in the studies reported by Salthouse (1991a). The present study was not primarily designed to investigate this issue, but some of the results are obviously relevant. For example, an analysis conducted with the composite-working-memory measure as the criterion revealed that the attenuation of the age-related effects on that measure was nearly 75%, from an R^2 of .251 for age when considered alone to a value of .063 after control of perceptual speed. Furthermore, the results summarized in Table 2 reveal that statistical control of the perceptual-speed measures reduced the age differences in reasoning accuracy by an amount larger than that produced by control of working memory and that there was little additional attenuation of the age differences by also controlling working memory.

In view of the apparent importance of perceptual speed in mediating the relations between age and performance in both reasoning and working-memory tasks, it is desirable to consider exactly what is measured by the tests of perceptual speed. Of particular concern is the possibility that because the digit-symbol substitution test involves nine digit-symbol pairs, it might represent memory factors as much as perceptual-speed factors. The issue of what is responsible for age-related differences in the digit-symbol test was recently investigated by Salthouse (1992). The following results, either based on original data or cited from previously published studies reviewed in that article, led to the conclusion that memory factors were relatively unimportant determinants of the age-related differences in digit-symbol performance: (1) age differences are still evident when all participants have learned the digit-symbol associations to a criterion of perfect recall; (2) age differences either remain constant or increase with additional opportunities to learn the digit-symbol pairs; (3) young and old adults have similar serial-position functions when response times to individual items are analyzed according to the position of the digit-symbol pair in the code table, suggesting that the code table was searched in the same manner by both age groups; (4) the age differences remain constant in relative terms as the number of digit-symbol pairs, and hence the presumed memory demands, is varied; (5) young and old adults devote the same proportion of their response time to merely copying the symbols; and (6) adding working memory to the prediction equation after perceptual speed resulted in little further attenuation of the age differences in digit-symbol performance. Results supporting an interpretation that the age differences were largely determined by the speed at which elementary operations could be executed were (1) high correlations between digit-symbol performance and performance on other measures of perceptual speed and (2) little unique age-related variance in digit-symbol performance after statistical control of variance in other measures of perceptual speed.

The case for the current composite measure of perceptual speed as a reflection of speed factors more than memory factors is even stronger than that for the digit-symbol measure alone because of the inclusion of the digit-digit measure in the perceptual-speed composite. It is difficult to imagine how memory factors could have contributed to performance on the digit-digit task in which *same/different* decisions were made about pairs of simultaneously presented digits. Of course, other influences may be operating, but it seems reasonable to suggest on the basis of the preceding arguments that the perceptual-speed composite used in this study probably does reflect some fairly basic aspect of the speed at which certain kinds of information can be processed.

One possible interpretation of the relation between speed and working memory is that working memory has a dynamic quality, perhaps somewhat analogous to someone trying to juggle several objects simultaneously. That is, just as the number of items that can be successfully juggled depends on the rate at which they can be caught and tossed, so might the limits on the number of distinct ideas that can be kept active (or mentally juggled) in working memory be set by the rate at which information can be processed. From this perspective, therefore, working memory might be interpreted as the set of items currently active in consciousness, and age differences in working memory might hypothesized to originate because increased age is associated with a reduction either in the ability to activate new information or in the ability to maintain the activation of old information.

The hypothesis that age-related differences in working memory might be mediated by reductions in the speed of executing relevant operations obviously needs to be confirmed with additional evidence from converging procedures. The cause of age differences in processing speed must also eventually be explained. Nevertheless, results from this and other recent studies seem to suggest a plausible, and testable, interpretation of the causes of adult age differences in cognition that merits further investigation.

REFERENCES

BADDELEY, A. D., & HITCH, G. J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 647-667), New York: Academic Press.

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

LIGHT, L. L. (1991). Memory and aging: Four hypotheses in search of data. *Annual Review of Psychology*, **42**, 333-376.

SALTHOUSE, T. A. (1990). Working memory as a processing resource in cognitive aging. *Developmental Review*, **10**, 101-124.

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

SALTHOUSE, T. A. (1991b). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Erlbaum.

SALTHOUSE, T. A. (1992). What do adult age differences in the Digit Symbol Substitution Test reflect? *Journal of Gerontology: Psychological Sciences*, **47**, 121-128.

SALTHOUSE, T. A., & BABCOCK, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, **27**, 763-776.

SALTHOUSE, T. A., LEGG, S. , PALMON, R., & MITCHELL, D. R. D. (1990). Memory factors in age-related differences in simple reasoning. *Psychology and Aging*, **5**, 9-15.

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, & 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**, 110-120.

STINE, E. A. L., & WINGFIELD, A. (1990). How much do working memory deficits contribute to age differences in discourse memory? *European Journal of Cognitive Psychology*, **2**, 289-304.

WECHSLER, D. (1981). *The Wechsler Adult Intelligence Scale–Revised*. New York: Psychological Corporation.

NOTES

1. Interactions involving the group variable with each predictor variable were also examined to determine whether the influence of the predictor varied across groups. None of these interactions was significant at the designated (.01) significance level.

2. Note that there are six possible premise types because a single relevant premise can appear in the first position when one, two, or three premises are presented, in the second position when two or three premises are presented, and in the third position only when three premises are presented.

3. An age group × trial type (one-relevant reasoning, recognition probe) ANOVA on the data from the restricted samples revealed that the age group × trial type interaction failed to reach the criterion significance level [$F(1,40) = 4.08$, $MS_e = 85.78$, $p = .05$].

(Manuscript received May 28, 1991;
revision accepted for publication January 7, 1992.)