Journal of Experimental Psychology:
Learning, Memory, and Cognition
1989, Vol. 15, No. 3, 507–516

0278-7393/89/$00.75

Effects of Adult Age and Working Memory on Reasoning and Spatial Abilities

Timothy A. Salthouse, Debora R. D. Mitchell, Eric Skovronek, and Renee L. Babcock
Georgia Institute of Technology

Three predictions were derived from the hypothesis that adult age differences in certain measures of cognitive functioning are attributable to age-related reductions in a processing resource such as working-memory capacity. Each prediction received at least some degree of empirical support in a study involving 120 males ranging between 20 and 79 years of age. First, older adults exhibited greater impairments of performance than did young adults when task complexity increased and more demands were placed on the limited processing resources; second, the magnitudes of these complexity effects were highly correlated across verbal (reasoning) and spatial (paper folding) tasks. Finally, statistical control of an index of a working-memory processing resource attenuated the effects of age on the measures of cognitive performance. It was concluded that further progress in understanding the mechanisms of the relation between age and cognitive functioning will require improved conceptualizations of the nature of working memory or other hypothesized mediating constructs.

One of the most frequently mentioned "interpretations" of adult age differences in fluid measures (cf. Cattell, 1971; Horn, 1982) or Type A (Hebb, 1942) measures of cognitive functioning attributes them to an age-related reduction in some type of general-purpose processing resources (see Salthouse, 1988a, 1988b, for review and discussion). In its simplest form, the resource argument consists of two assumptions and a conclusion: (a) the assumption that processing resources are required for many, but not all, cognitive processes; (b) the assumption that, for largely unspecified reasons, increased age in adulthood is associated with a diminished supply of available processing resources; and (c) the conclusion that the age-related reduction in the quantity of processing resources results in poorer performance in tasks containing resource-demanding processes. Unfortunately, although the frequent reference to processing resources in discussions of cognitive aging phenomena suggests that many researchers find this general argument compelling, the processing-resources interpretation is severely weakened by two conceptual problems—vagueness of the fundamental construct and ambiguity of the relevant mechanisms.

The first problem is that the specific nature of the key concept in this category of interpretation—processing resources—has seldom been discussed. Instead, researchers have employed a variety of synonyms such as *effort*, *energy*, or *capacity* without ever specifying exactly what is meant by these terms. Adjectives are occasionally added with the apparent intent of increasing the precision of the terms, for example, *cognitive effort*, *mental energy*, *attentional capacity*, and *working-memory capacity*, but these elaborations have typically not been accompanied by more explicit descriptions that would remove the ambiguity inherent in the resources construct.

A second major problem with past usages of the concept of processing resources is that there has been little or no attempt to specify the mechanisms by which processing resources might influence cognitive performance. The primary question in this context, which has been ignored in almost all previous references to the resources construct within the cognitive aging literature, is how a limited supply of some entity contributes to lower performance in tasks of memory, reasoning, and spatial abilities. Furthermore, unless there is at least some empirical evidence supporting the hypothesized causal relation, it is difficult to view speculations containing references to processing resources as serious scientific hypotheses.

The goal of the present research was to address the aforementioned problems while investigating the hypothesis that at least some of the adult age differences in certain cognitive tasks are mediated by age-related declines in a type of processing resource. In the following section the reasoning underlying predictions derived from the resource perspective is discussed; the criteria used in selecting tasks to test those predictions are described in subsequent sections.

Predictions From the Processing-Resource Perspective

Two initial predictions can be derived from the resources perspective based on the age-complexity phenomenon, that is, the tendency for the magnitude of the age differences in performance to increase with the hypothesized complexity of the task (see Salthouse, 1985, pp. 183–190, for review and discussion). The first prediction is that qualitatively similar age-complexity patterns should be evident across different cognitive tasks; the second is that the magnitude of the complexity effects in different tasks should be comparable for a given individual. The key to both predictions is the assumption that the performance decline associated with an increase in task complexity is at least partially attributable to greater

---

This research was supported by National Institute on Aging Research Grants KO4AG00372 and RO1AG06826 to the senior author.

Correspondence concerning this article should be addressed to Timothy A. Salthouse, School of Psychology, Georgia Institute of Technology, Atlanta, Georgia 30332.

demands on a finite quantity of processing resources. If older adults have smaller supplies of the relevant processing resources than do young adults, then they would be expected to suffer greater performance impairments as the demands on their more limited resources increase. Moreover, if a common processing resource is involved in the complexity-related performance decrements in different cognitive tasks, then the quantity of resources available to an individual should influence the magnitude of the complexity effects in each task, and consequently, those effects should be significantly correlated across tasks.

Another expectation from the view that a diminished supply of general-purpose processing resources contributes to the age differences in different cognitive tasks is that statistical control of an index of the hypothesized processing resources should attenuate the magnitude of the age differences in measures of cognitive performance. The degree of attenuation will naturally vary across measures, depending on the importance of the relevant processing resource to the age differences in particular tasks. Moreover, it is probably unrealistic ever to expect complete elimination of age differences by statistical control of an index of processing resources because there are likely to be some nonresource determinants of performance (e.g., sensory acuity, quantity or quality of relevant knowledge, etc.) in most cognitive tasks. Nevertheless, if it is true that age differences in measures of cognitive performance are at least partially determined by age differences in some type of processing resource, then statistical control of an index of the latter should result in the attenuation of the effects of age on the former.

Selection of a Resource Index

Examination of the third prediction discussed above obviously requires the availability of an index of the individual's quantity of processing resources. Two major criteria were employed in the current project to guide in the selection of this index: demonstrable reliability and clear theoretical relevance to the resources construct. Reliability is essential because the measure must be stable and consistent if it is to reflect an enduring traitlike characteristic such as the quantity of one's processing resources. Some indication of the reliability of the measure is also necessary to allow meaningful interpretation of the correlational results because the range of possible correlations between one variable and another is obviously limited by the magnitude of the correlations the variables have with themselves, that is, by their reliabilities.

There are two aspects to the theoretical relevance criterion. The first is that processing resources should be positively related to various measures of cognitive performance, and the second is that they should be negatively related to adult age. Stated somewhat differently, if processing resources truly mediate some of the age-related declines observed in cognitive functioning, then the index of the quantity of processing resources should itself decline across the adult years, and there should be reason to believe that higher amounts of the resource are associated with better levels of cognitive performance.

A variety of potential measures appear to possess the desired characteristics, but determination of the mechanisms by which processing resources influence cognitive performance was considered most feasible with measures of the working-memory conceptualization of processing resources. Furthermore, working memory is often postulated to play a central role in a variety of cognitive tasks (e.g., Baddeley 1986; Case, 1985), and there have been many reports that aging is associated with an impairment in working memory (e.g., Craik & Rabinowitz, 1984) or active short-term memory (e.g., Welford, 1958).

Baddeley (1986), Baddeley, Logie, and Nimmo-Smith (1985), Case (1985), Daneman and Carpenter (1980), and others have all claimed that a critical aspect of working memory is that it involves the simultaneous storage and processing of information. An important feature of a task assessing working memory, therefore, is that it must require the maintenance of some information during the processing of that or other information.

The task selected to assess this type of working memory is a modification of the computational span task used by Salthouse (1988a) and Salthouse and Prill (1987). It is similar to the reading span task employed by Daneman and Carpenter (1980), in which the subject reads or listens to sentences while remembering the last word in each sentence, and the counting span task employed by Case, Kurland, and Goldberg (1982), in which the subject views slides containing a variable number of dots, and the number of dots in each slide must be counted while remembering the numbers from earlier slides. The computational span task consists of the subject's performing simple arithmetic problems while simultaneously remembering a designated digit from each problem. Earlier research (Salthouse, 1988a) has indicated that although the computational span did not have impressive reliability (i.e., estimated reliability coefficients of .65 for 20 young adults and .57 for 20 older adults), it was nevertheless moderately correlated with other measures presumed to reflect working memory (e.g., average correlations of .32 with a measure of the number of digits that can be correctly recalled in the reverse sequence of presentation, and .56 with a measure of the number of digits that can be correctly recalled after first subtracting the number 2 from each digit).

Selection of Cognitive Tasks

Several criteria were also considered in the selection of the cognitive tasks used to examine the predictions from the processing-resource hypothesis. One criterion was that the tasks should be similar to psychometric tests that have been considered to represent distinct domains of cognitive activity. The rationale for this criterion is to ensure that the hypothesized working-memory processing resource is truly general, not specific to a particular type of cognitive task. A second criterion was that the tasks should allow within-task variation of hypothesized resource demands by quantitative, not simply qualitative, manipulation of complexity. The requirement that complexity be varied quantitatively was introduced to minimize the possibility that complexity-related differences in performance could be attributed to new processing components that were added with qualitative changes in the task, rather than to an increase in the hypothesized resource demands. And finally, the third criterion considered in the selection of the cognitive tasks was that the tasks should be analyzable in order to identify the mechanisms by which processing resources influence cognitive performance. That is, if the processing resource to be investigated is working-memory capacity, then it should be possible to indicate how a smaller working-memory capacity might result in lower levels of performance in these tasks. Two tasks satisfying these criteria, and employed in the present project, are integrative verbal reasoning and spatial paper folding.

Sample trials in the four conditions of the integrative verbal-reasoning task are illustrated in Figure 1. Note that one to four premises describing a relation between two variables were presented, followed by a question concerning the status of one variable, given a specified change in another variable. Each premise and question was displayed successively after removal of the previous premise. Variables within the premises were selected randomly from the alphabet, and the premises, which always described a relation between adjacent letters in the alphabet, were presented in a random order to maximize demands on memory.

Average decision accuracy was expected to decrease as the number of premises in the problem increased because of the greater demands on the limited working-memory capacity. That is, it was assumed that successively larger requirements on working memory are imposed with each additional premise because more earlier premises must be maintained during the registration and encoding of later premises.

The task was also designed to allow a direct assessment of the role of working-memory factors in the complexity-related performance decrements. This was accomplished by examining the effect on decision accuracy of the number of premises presented when only one of those premises was actually relevant to the decision. These "one-relevant trials" are all similar in that the same type of decision is required, involving a question about the status of one variable when the relation between that variable and the causal variable had been described in a single premise, and only the context in which the relevant information is presented changes. For example, if the question in the four-premise condition illustrated in Figure 1 was about B and C, then only one premise (the second) would be relevant to the decision, despite the fact that a total of four premises had actually been presented.

1 Premise	F and G do the SAME. If F DECREASES, what will happen to G?
2 Premises	M and N do the OPPOSITE. L and M do the SAME. If L INCREASES, what will happen to N?
3 Premises	R and S do the OPPOSITE. T and U do the SAME. S and T do the OPPOSITE. If R DECREASES, what will happen to U?
4 Premises	D and E do the SAME. B and C do the OPPOSITE. C and D do the SAME. E and F do the OPPOSITE. If B INCREASES, what will happen to F?

Figure 1. Illustration of the sequence of displays for the four conditions in the integrative-reasoning task.

The decision processes can be assumed to remain constant across conditions involving different numbers of premises when only one of those premises is relevant to the decision. Therefore, a reduction in decision accuracy when additional premises are presented can presumably be attributed to a loss of necessary information from some type of working memory. To the extent that the slope of the function relating decision accuracy to number of presented premises in one-relevant trials is similar to that based on the data from all trials, one could infer that reduced accuracy with additional premises is largely caused by a failure to preserve early information during the presentation and processing of later information.

Sample trials in the four conditions of the spatial paper-folding task are illustrated in Figure 2. Successive displays in this task represented a square piece of paper folded from one to four times, the punching of a hole in the folded paper, and a pattern of circles indicating the locations of the punched holes in the unfolded paper. Although Figure 2 illustrates only the outcome of each fold, implementation of the experimental task on a microcomputer allowed dynamic displays of the folds as they were in progress, rather than simply portraying static representations of the outcome of each fold. The task for the subject was to decide whether the pattern of holes in the final display was consistent with the pattern that would result from the earlier sequence of folds and punch location.

As in the reasoning task, decision accuracy was expected to decrease as the number of relevant pieces of information was increased. The relevant information in this task was presumably the type (e.g., horizontal, vertical, diagonal) and location (e.g., top, middle, bottom, left, center, right) of each fold, and progressively more of this information had to be represented as the number of folds in the problem increased.

Because the paper-folding and integrative-reasoning tasks were designed to be structurally equivalent, direct analyses of the role of working memory in the paper-folding task were possible in a manner analogous to that in the reasoning task. That is, memory factors would be implicated if similar effects of the number of presented folds on decision accuracy were evident across all trials and on trials when only one fold was relevant to the decision. For example, if a trial in the four-fold condition consisted of nonoverlapping folds of each corner, as portrayed in the first two folds in the example of a four-fold trial in Figure 2, then only one of the folds would be relevant to the decision, regardless of the location of the punched hole. A discovery of parallel functions relating the number of presented folds to decision accuracy in data from all trials and in data from only one-relevant trials would therefore lead to the inference that much of the complexity-related performance reduction was attributable to a loss of the relevant information.

Review of Predictions

We will now briefly review the predictions derived from the hypothesis that at least some of the age-related differences

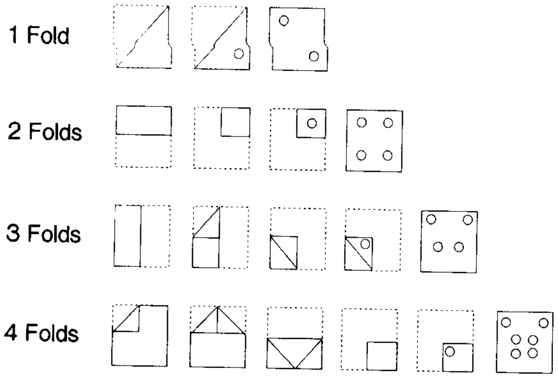

Figure 2. Illustration of the sequence of displays for the four conditions in the paper-folding task.

in certain cognitive tasks are attributable to age-related declines in a working-memory processing resource. First, it is predicted that the effects of the number of premises or the number of folds should be greater with increased age because older adults are assumed to have a smaller supply of available resources than young adults to cope with greater demands on those resources. The second prediction is that the magnitude of the complexity effects in the integrative-reasoning and paper-folding tasks should be correlated with one another if a common processing resource contributes to both effects. And finally, it is predicted that statistical control of the computational span working-memory variable should significantly attenuate the effects of age on the measures of accuracy in the reasoning and paper-folding tasks because age effects in the latter variables are presumed to be mediated by age effects in the resource construct indexed by the former variable.

In addition to examining these resource predictions, the current project was also designed to allow a decomposition of the reduction in performance associated with an increase in the number of premises or folds. That is, if the decline in accuracy associated with an increase in task complexity is primarily attributable to the loss of relevant early information during the processing of later information, then comparable performance reductions should be evident across all trials and in trials when only one fold or premise is relevant to the decision. On the other hand, if processes other than the simultaneous retention and processing of information are involved in the complexity-related performance declines, then the effects of number of premises or number of folds should be much smaller in trials when only a single premise or fold is relevant to the decision.

Method

Subjects

A total of 120 males, 20 from each decade from the 20s through the 70s, participated in the project. All but a few, who were still students, were alumni of Georgia Institute of Technology, an institution with a primarily engineering and technically oriented curriculum. The sample can therefore be considered relatively homogeneous with respect to intellectual level, socioeconomic class, and educational experiences, although there was moderate variation in current occupation. The mean years of education reported was 17.1, and this variable was not correlated with chronological age (i.e., $r = -.01$). Self-reported health on a five-point scale (1 = *excellent*, 5 = *poor*) averaged 1.4, and 98% of the individuals reported themselves to be in at least average health. The self-reported health rating correlated .26 ($p < .01$) with chronological age, indicating slightly poorer evaluation of one's health with increased age. This variable was introduced as a covariate in all of the data analyses reported below, but it did not alter any of the results and accounted for less than 1% of the variance in all of the analyses; thus it is not discussed further.

Procedure

The original experimental design was for each subject to perform the following tasks: the computational span task, four blocks of 40 trials each on the integrative-reasoning task, four blocks of 40 trials each on the paper-folding task, and then a repetition of the computational span task. However, in order to ensure at least a minimum amount of data from each subject in each task, a schedule was followed in which the subject was switched to the next task after a specified time had elapsed from the beginning of the session. This proved propitious because a moderate percentage of the subjects were unable to complete the desired number of trial blocks within the time limits of 2.5 hr imposed on the session. Results are therefore reported from the first two blocks of the reasoning and paper-folding tasks completed by everyone, and the data from subjects completing four blocks are used to examine the reliability of the cognitive performance measures. (It should be noted, however, that the same pattern of results as that reported was also evident when the analyses were conducted on all of the data, with the number of trial blocks completed in each task as a covariate.)

Figure 3A illustrates the implementation of the computational span task in the current project. Each trial consisted of a series of successively presented arithmetic problems, with the second digit in each problem highlighted as a to-be-remembered digit. Subjects were instructed to answer each arithmetic problem as it was displayed and then to recall the sequence of target digits when the word *recall* appeared. In the four-problem example illustrated in Figure 3A, the subject should answer *5*, *3*, *4*, and *6* to the arithmetic problems, followed by *2-4-2-1* in response to the *recall* probe. Target digits were randomly selected from the set 1–9, with the constraints that the target digit was not identical to the correct answer on that problem and that successive to-be-remembered digits were not identical.

Figure 3B illustrates that a double random staircase psychophysical procedure was used to identify each subject's computational span. That is, two independent sequences of trials were presented in random alternation within the same block of trials. The number of arithmetic problems in each sequence was increased if the target items were recalled correctly, and it was decreased if they were recalled incorrectly. No change in the number of arithmetic problems for a given sequence was introduced if an error was made in any of the component arithmetic problems for that trial. The task was terminated when the two sequences had problem lengths within two of each other for six consecutive trials or when a total of 25 trials had been presented. The span was the average of the values from the two sequences at the point of termination.

The data in Figure 3B are from a representative administration of the task. Note that the task started with a nine-problem trial in Sequence A, but not all nine digits were correctly recalled; thus the next trial in that sequence had only eight problems. The recall attempt was also unsuccessful on this trial, and so the next trial in the sequence

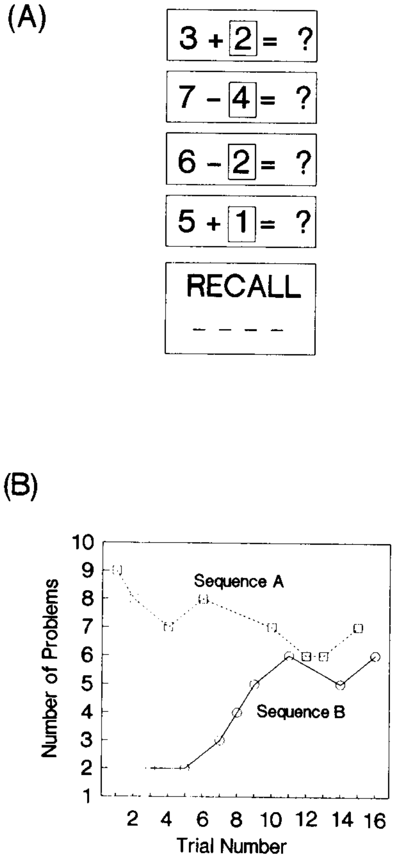

Figure 3. Panel A: Illustration of the sequence of displays in the computational span task. Panel B: Example of how the span is estimated from the convergence of two independent sequences of trials.

contained only seven problems. However, the seven-problem trial was not presented until Trial 4 because there was a switch to Sequence B for Trial 3. This sequence started with a two-problem trial, but an arithmetic error was committed on that trial; thus the next trial in this sequence, Trial 5, also contained only two arithmetic problems. Eventually the two sequences converged to the specified criterion, yielding an estimated computational span of six items.

All phases of the task were self-paced in that the next arithmetic problem within a trial was not presented until an answer had been entered on the computer keyboard to the previous problem, and the next trial was not presented until the appropriate number of digits had been entered during the recall phase for that trial.

The reasoning and paper-folding tasks were presented in blocks of 40 trials. Each block contained 10 trial types consisting of n = 1, 2, 3, or 4 premises or folds, with from 1 to n premises or folds relevant to the decision. In other words, there were four trial types containing four premises or folds (with 1, 2, 3, or 4 relevant premises or folds), three trial types containing three premises or folds (with 1, 2, or 3 relevant premises or folds), and so forth. Each block contained a random arrangement of two positive (*increase* or *yes*) and two negative (*decrease* or *no*) trials of each type. Feedback indicating the correct answer in the trial was displayed after each response.

A trial was initiated in both tasks by pressing the *enter* key on the computer keyboard. The first fold or premise was then displayed, and each successive premise or fold was displayed by pressing *enter* again. The question display in the reasoning task was accompanied by the words *increase*, on the lower left of the screen, and *decrease*, on the lower right of the screen; decisions were communicated by pressing the "Z" or "/" key on the keyboard, respectively. In the paper-folding task, the display of the pattern of holes was accompanied by the words *no* on the lower left of the screen and *yes* on the lower right of the screen; decisions were communicated by pressing the "Z" or "/" key, respectively.

Results

General Age Effects

Initial analyses focused on the variables of mean computational span (i.e., the average of the two spans for a given individual, designated CSpan) and mean percentage correct across all 80 trials in the first two trial blocks for both the reasoning and paper-folding tasks. Averages of these variables across the 20 subjects in each age decade are displayed in Figure 4. The most striking feature of this figure is that the age trends are remarkably similar across the three variables. In fact, the slopes of the linear regression equations relating age to performance were −.43%/year for both reasoning and paper folding, and −.04 digits/year for CSpan.

Table 1 contains the matrix of correlations among the primary variables, with estimated reliabilities in parentheses along the diagonals. Reliability of the CSpan measure was obtained by boosting the correlation between the values from the first and second computation span assessments by the Spearman-Brown formula to estimate the reliability of the average value. The reasoning and paper-folding reliabilities are correlations between the values from the first two trial blocks and the values from the last two trial blocks for the 49 subjects with four blocks in the reasoning task and for the 55 subjects with four blocks in the paper-folding task.

The data summarized in Figure 1 and Table 1 were subjected to hierarchical regression analyses to examine the effects of age on reasoning and paper-folding performance after partialing out the variance associated with the CSpan measure. For the reasoning task, the R^2 attributable to age when it was considered alone was .278, $F(1, 118) = 45.51$, $MS_e = 136.48$, $p < .01$, and this was reduced to .119, $F(1, 117) = 21.41$, $MS_e = 123.95$, $p < .01$, after removing the variance associated with the CSpan working-memory index. Similar

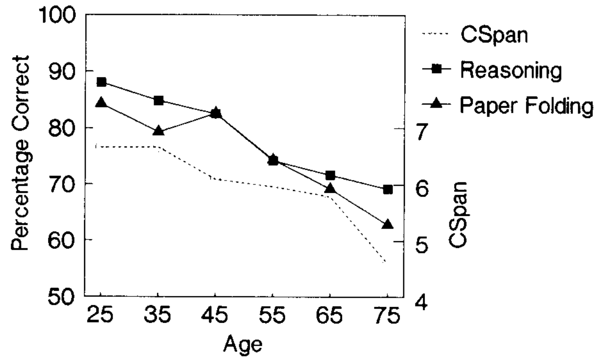

Figure 4. Mean values of computational span task (CSpan) and average percentage correct in the reasoning and paper-folding tasks for the 20 individuals in each age decade.

Table 1
Correlations Among Primary Variables for all 120 Subjects

Variable	1	2	3	4
1. Age	X	−.46	−.53	−.53
2. CSpan		(.78)	.48	.38
3. Reasoning			(.86)	.62
4. Paper folding				(.86)
M	49.3	5.94	78.4	75.4
SD	16.8	1.37	13.7	13.7

Note. All correlations are significant at $p < .01$.
CSpan = computational span. Estimated reliabilities are in parentheses.

results were evident in the paper-folding task because the R^2 for age, when considered alone, was .282, $F(1, 118) = 46.45$, $MS_e = 136.59$, $p < .01$, and this was reduced to .161, $F(1, 117) = 27.08$, $MS_e = 133.20$, $p < .01$, after statistical control of the CSpan measure of working-memory capacity.

Quantitative estimates of the relative importance of the CSpan-indexed working-memory factor to the age differences in the verbal integrative-reasoning and spatial paper-folding tasks can be derived from the results just described. That is, the proportional reduction in age-associated variance achieved by statistical control of the CSpan index can be determined by dividing the difference in age-associated variance with and without the control of CSpan by the total amount of age-related variance. These computations revealed that statistical control of the CSpan index of working-memory capacity attenuated the age effects in the reasoning task by 57.2%, and it attenuated the age effects in the paper-folding task by 42.9%. Although the absolute proportions differ across the two tasks, the results are consistent in suggesting that a moderate proportion of the age effects in both reasoning and paper-folding performance is mediated through age-related reductions in working-memory capacity, as the latter is indexed by the CSpan variable.

Practice Effects

In all three tasks, practice resulted in better performance, but it did not substantially alter the magnitude of the age relations. That is, the first computational span assessment yielded an average value of 5.65, with an age correlation of −.39, while the second assessment resulted in an average of 6.21, with an age correlation of −.45. Practice effects in the reasoning and paper-folding tasks were examined for the subjects completing four trial blocks in each task. The 49 subjects with four blocks in the reasoning task averaged 80.3% correct decisions for the first two blocks (age correlation = −.58) and averaged 82.8% for the last two blocks (age correlation = −.46). The 55 subjects with four blocks in the paper-folding task averaged 78.1% in the first two blocks (age correlation = −.44) and 82.0% in the second two blocks (age correlation = −.55).

Integrative-Reasoning Performance

Decision accuracy in the integrative-reasoning task was computed as a function of the number of premises presented across all trials and across trials when only one premise was relevant to the decision. Means across all 120 subjects of these values are illustrated in Figure 5. As expected, increasing the number of premises presented, and thus presumably increasing the demands on the limited working-memory capacity, resulted in progressively lower levels of average accuracy. Of particular interest, however, is the similarity in the functions for all trials (i.e., a slope of −8.8% per premise) and for trials when only one premise was relevant to the decision (i.e., a slope of −7.4% per premise). On the basis of the argument outlined earlier, these findings can be interpreted as suggesting that much of the performance decline associated with increased complexity in this task is attributable to a failure to preserve relevant information in the more complex conditions.

Analyses of the effects of age on these complexity slopes were restricted to subjects with an accuracy of at least 85% in the one-premise condition in order to minimize the possibility of floor effects in the slope measures. That is, low performance in the simplest condition precludes meaningful estimates of the complexity slopes because of insufficient range of performance variation from the one-premise to the four-premise condition. The value of 85% was somewhat arbitrary; it was high enough to allow an adequate range for lower accuracy when more premises were presented, but yet not too high to exclude a large proportion of subjects. The correlation between age and accuracy in the one-premise condition was −.30, and thus more old than young adults were excluded with this procedure. The restricted sample of 104 subjects is still representative of the larger sample, however, because the correlation between age and average percentage correct was −.48 in this group compared with −.53 in the total sample. The mean age in the restricted sample was 46.9 years (*SD* = 16.4) compared with 49.3 years (*SD* = 16.8) for the entire sample.

Reliabilities of the complexity slope parameters relating decision accuracy to number of presented premises were estimated by determining the correlation between the slopes from the first two trial blocks and from the last two trial blocks for subjects with accuracy of at least 85% in the one-premise condition in both sets of trials. The correlations for the 41 subjects with the appropriate data were .76 for the

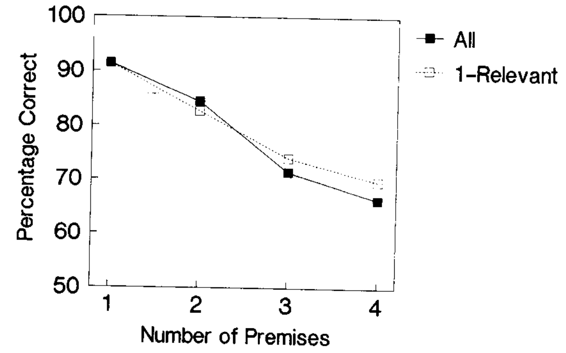

Figure 5. Mean percentage correct as a function of number of premises presented for all trials and for trials with only one relevant premise in the reasoning task.

slope based on all trials, and .56 for the slope based on the one-relevant trials.

Mean slopes for the 13–20 subjects in each age decade are illustrated in Figure 6. Despite the apparent floor effect beginning in the decade of the 50s, the age correlations were −.46 for the slope based on all trials and −.42 for the slope based on only one-relevant trials. These very similar patterns suggest that the composition of the complexity effect is relatively constant across the adult years. Stated somewhat differently, the fact that nearly parallel age trends were evident in the complexity slopes from all trials and from only one-relevant trials can be interpreted as indicating that most of the age-related increase in the effects of additional premises is attributable to age-related increases in the loss of relevant information, that is, to working-memory limitations.

Results from the hierarchical regression analyses with the slope based on all trials were that the R^2 attributable to age when it was considered alone was .212, $F(1, 102) = 27.37$, $MS_e = 25.31$, $p < .01$, and the increment in R^2 associated with age after controlling for CSpan was .148, $F(1, 101) = 19.20$, $MS_e = 25.23$, $p < .01$. Corresponding results for the slopes based on one-relevant trials were that the R^2 attributable to age by itself was .174, $F(1, 102) = 21.46$, $MS_e = 38.11$, $p < .01$, and the increment in R^2 associated with age after controlling for CSpan was .141, $F(1, 101) = 17.36$, $MS_e = 38.47$, $p < .01$. Expressed as measures of percentage of attenuation due to statistical control of the CSpan index, the values were 30.2% for the slope from all trials and 19.0% for the slope from one-relevant trials.

Paper-Folding Performance

Analyses similar to those conducted on the measures of reasoning performance were conducted on the measures of paper-folding performance. Mean decision accuracy as a function of the number of folds presented is displayed in Figure 7 for all trials and for trials with only one relevant fold. Regression slopes based on the average data were −6.9% per fold with all trials and −5.3% per fold with one-relevant trials. Figure 7 reveals that accuracy with one-relevant fold in the three-fold condition appears abnormally high, and it was suspected that this might be due to the existence of a high proportion of either simple folds or obvious distractors in this condition, resulting in a set of particularly easy problems. However, ratings from four judges of the perceived difficulty of the one-relevant problems increased monotonically with additional folds and did not correspond to the observed performance pattern. We are therefore unable to offer a reason for the nonmonotonic trend in the one-relevant data.

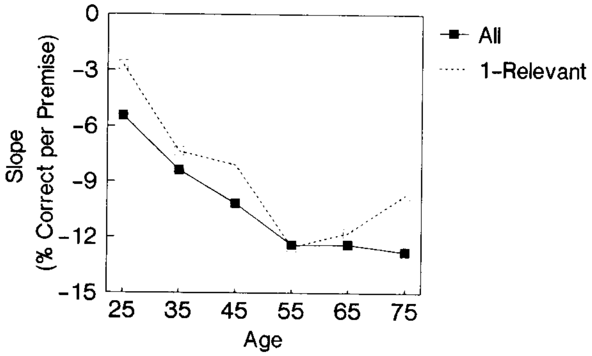

Figure 6. Slopes of the functions relating percentage correct to number of premises presented for all trials and for trials with only one relevant premise for the 13–20 individuals in each age decade.

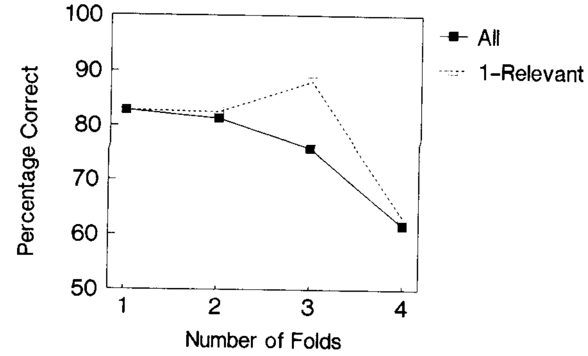

Figure 7. Mean percentage correct as a function of number of folds presented for all trials and for trials with only one relevant fold in the paper folding task.

Age effects in the complexity slopes were examined after first excluding subjects with accuracies of less than 85% in the one-fold condition. The rationale for this restriction is the same as that in the analyses of the reasoning data. The correlation of −.32 between age and percentage correct in the one-fold condition indicates that a larger number of old than young adults were eliminated with this procedure, but the restricted sample is still representative of the complete sample because the correlation between age and average percentage correct was −.53 in both samples. The mean age of the 78 subjects in the restricted sample was 45.3 (*SD* = 15.3), compared with the mean of 49.3 (*SD* = 16.8) for all 120 subjects.

Reliabilities of the slope parameters were estimated from the correlations between the values from the first two trial blocks and the last two trial blocks for the 36 subjects with average accuracies in the one-fold condition of at least 85% in both sets of trials. The correlations were .42 for the slopes from all trials and .39 for the slopes from one-relevant trials.

Mean slopes for the 7–17 subjects in each age decade are displayed in Figure 8. The similar age trends for the two slopes, together with the nearly identical age correlations (i.e., $r = -.47$ for slopes from all trials, $r = -.45$ for slopes from one-relevant trials), suggest that much of the age-related increase in the effects of task complexity is attributable to a loss of relevant information from working memory.

Hierarchical regression analyses revealed that the R^2 in the complexity slope from all trials attributable to age when age was considered alone was .218, $F(1, 76) = 21.15$, $MS_e = 15.88$, $p < .01$, and the increment in R^2 associated with age after the control of CSpan was .177, $F(1, 75) = 17.34$, $MS_e = 15.77$, $p < .01$. Corresponding values for the slope from one-relevant trials were .199 for R^2 due to age alone, $F(1, 76) = 18.89$, $MS_e = 41.76$, $p < .01$, and .158, $F(1, 75) = 15.19$, $MS_e = 41.31$, $p < .01$, for the increment in R^2 due to age after control of CSpan. The percentage of attenuation of the age effects by control of the CSpan index of working-memory

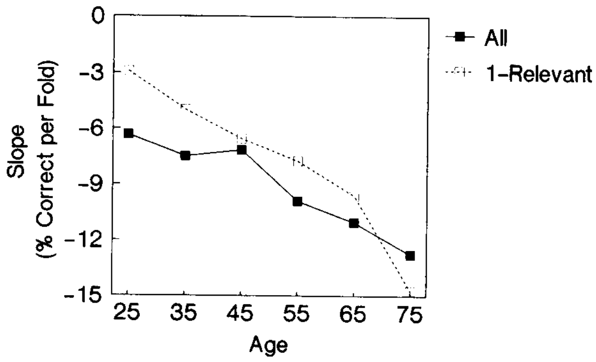

Figure 8. Slopes of the functions relating percentage correct to number of folds presented for all trials and for trials with only one relevant fold for the 7–17 individuals in each age decade.

capacity was therefore 18.8% for the slope from all trials and 20.6% for the slope from one-relevant trials.

Interrelations of Performance Measures

Correlations among all relevant measures for the 73 individuals with accuracy of at least 85% in both the one-premise reasoning condition and the one-fold paper folding condition are presented in Table 2. Of particular interest in this table are the correlations of .63 between the complexity slopes based on all trials and .45 between the complexity slopes based on one-relevant trials. These values are impressive because in both cases they are close to the average of the estimated reliabilities of each measure (i.e., .76 and .42 for the slopes from all trials, and .56 and .39 for the slopes from one-relevant trials). It therefore appears that nearly all of the reliable variance in the complexity slope measures from the two tasks is shared, or common, variance.

Discussion

Three predictions, derived from the view that one determinant of age differences in certain reasoning and spatial ability cognitive tasks is an age-related reduction in a working-memory processing resource, were examined in this study. The results provide strong support for two of the predictions and moderate support for the third prediction.

One prediction was that if cognitive performance declined with increased task complexity because of greater demands on a limited processing resource, then the magnitude of those performance declines should increase with increased age because of the hypothesized age-related reduction in processing resources. The results summarized in Figures 6 and 8 indicate that this prediction was clearly confirmed. Furthermore, the discovery that very similar age trends were evident in the complexity slopes derived from trials in which only a single premise or fold was relevant to the decision suggests that most of these age effects can be attributed to a failure to retain information from early premises or folds during the presentation of subsequent premises or folds. In this respect, limitations of the working-memory processing resource appear to be implicated as a major determinant of the age-complexity effects in the present study.

A second prediction examined in the study was that if the complexity-related performance declines were attributable to limited quantities of a relevant processing resource, then the magnitude of those declines for a given individual should be similar across different tasks involving that same resource. The complexity manipulations in the current reasoning and paper-folding tasks were assumed to involve a resource related to working-memory capacity, and thus it was expected that the slopes of the functions relating accuracy to complexity in each task would be significantly correlated with one another. This expectation was confirmed in that the correlation between the two slopes based on all trials was .63, and the correlation between the slopes from one-relevant trials was .45. Because the correlation with the other measure was nearly as high as the expected correlation of the measure with itself (i.e., the average of the estimated reliabilities was .59 for the slopes from all trials and .46 for the slopes from one-relevant trials), it can be concluded that the two measures share nearly all of their reliable variance and, hence, probably reflect the same construct.

Table 2
Correlations Among Select Variables for 73 Subjects With at Least 85% Correct in the One-Premise Reasoning Condition and in the One-Fold Paper Folding Condition

Variable	1	2	3	4	5	6	7	8
1. Age	X	−.26	−.39	−.53	−.43	−.39	−.45	−.44
2. CSpan		(.63)	.35	.33	.28	.13	.18	.20
3. R%C			(.86)	.71	.85	.62	.64	.50
4. PF%C				(.86)	.63	.39	.67	.53
5. RS					(.76)	.78	.63	.56
6. R1S						(.56)	.47	.45
7. PFS							(.42)	.77
8. PF1S								(.39)
Mean	44.6	6.33	83.6	82.9	−9.57	−8.05	−8.15	−6.36
SD	15.3	1.12	10.2	8.1	5.83	6.52	4.24	7.03

Note. CSpan is computational span; R%C is average percentage correct in the reasoning task; PF%C is average percentage correct in the paper-folding task; RS and R1S are complexity slopes for all and one-relevant trials in the reasoning task; and PFS and PF1S are complexity slopes for all trials and for one-relevant trials in the paper-folding task.
Correlations greater than .30 are significant at $p < .01$. Estimated reliabilities are in parentheses.

Although the results relevant to the first two predictions suggest that working-memory factors are involved in the tendency for age differences to increase as task complexity increases, significant age effects were also found in the least complex (i.e., one premise and one fold) conditions. A plausible inference from this combination of findings is that at least two factors contribute to the age differences observed in the current integrative-reasoning and paper-folding tasks: limitations of working memory and some as-yet-unspecified factors that are responsible for the differences observed when the tasks appear to place little or no demands on working memory.

Results from the analyses of the prediction that statistical control of an index of processing resources will attenuate the age differences in measures of cognitive performance also seem consistent with this dual-determinant interpretation. That is, although partialing out the variance associated with the CSpan measure of working memory attenuated the age effects in the measures of overall percentage correct in the reasoning and paper-folding tasks, it did not completely eliminate them.

These results are subject to at least two quite distinct interpretations. On the one hand, it could be argued that the discovery that at least some of the age differences in the present reasoning and spatial tasks appear to be mediated by age-related reductions in working memory, as the latter is indexed by the CSpan measure, represents an important step toward understanding why age differences occur in certain cognitive tasks. According to this perspective, there are almost certainly multiple determinants contributing to the age-related differences in cognitive functioning, and consequently it would have been unrealistic to have expected complete elimination of the age differences by statistical control of any single variable. The fact that an appreciable proportion of the age differences can apparently be attributed to a factor related to working memory might therefore be viewed as an impressive outcome.

On the other hand, it could also be argued that the failure to eliminate, or at least dramatically attenuate, the age differences after statistical control of the index of the working-memory processing resource is inconsistent with the presumed importance of the processing-resource construct. That is, the results of this study and of several related ones involving other combinations of cognitive measures and resource indexes (e.g., Salthouse, 1988a, 1988b; Salthouse, Kausler, & Saults, 1988) suggest that less than 50%, and perhaps only 15%–20%, of the overall age differences in fluid cognitive tasks can be accounted for by age-related reductions in processing resources as assessed with available measures. Because the assumption implicit in many discussions is that an age-related reduction in processing resources is a major determinant of the age differences observed in certain cognitive tasks, these results, which at best suggest a relatively weak influence of processing resources, might be viewed as rather disappointing.

Although debates could obviously continue concerning the validity of each of these interpretations, a more productive focus for future efforts might consist of carefully examining reasons why there was not greater attenuation of the age effects after statistical control of the measure of working memory capacity. Of course, one possibility is that the attenuation was small because working-memory factors play only a minor role in the age differences in the current measures of cognitive functioning. This view is clearly plausible with some variables, but it seems unlikely with the measures in the present tasks, particularly the slope measures representing the amount of change in accuracy with each additional premise or fold. That is, the nearly equivalent slopes for all trials and for one-relevant trials indicate that accuracy decreases with additional premises or folds in large part because of an inability to preserve relevant information during the processing of other information, a situation that appears to epitomize failure of working memory.

A second possible reason for the relatively small attenuation of the age effects by statistical control of the working-memory resource is that cognitive performance may not be a simple linear function of the quantity or capacity of working memory. For example, there may be something analogous to a resource threshold whereby cognitive performance is not impaired if the working-memory resource exceeds some minimum value, but it is severly impaired as the resource falls below the minimum or threshold value. Although the existence of a threshold would normally be expected to yield a distinctive step function, patterns of this type may be difficult to identify in group data if people vary in the positions of their respective thresholds. Therefore, unless it can be assumed that the relation between resource quantity and cognitive performance is reasonably linear and relatively similar across age groups, linear regression procedures may yield misleading estimates of the resource contribution to the age-related cognitive differences.

Another factor that may have contributed to the relatively modest attenuation of the age effects after statistical control of an index of the working-memory processing resource is the possibility that working memory was not adequately assessed by the CSpan measure. At least three issues can be raised concerning the validity of the CSpan measure as an index of working memory. One is related to the fact that even though several criteria were considered in selecting the CSpan measure, it still has all of the limitations associated with single-variable assessment of a complex construct (e.g., narrow and test-specific reflection of construct, restricted reliability, etc.). Had suitable alternative measures of working memory been available, these limitations could have been minimized by using multiple indicators of the working-memory construct.

A second potential weakness of the CSpan measure with respect to its validity as an index of working memory concerns the possibility that the measure may not have reflected the same characteristic in every subject. The CSpan measure was selected over alternative measures because it allows accurate monitoring of both processing (accuracy of arithmetic computations) and storage (accuracy of digit recall), and it has respectable reliability when assessed with appropriate procedures. However, observation of the subjects performing the computation span task revealed that at least some subjects attempted to rehearse the to-be-remembered digits during the time in which arithmetic problems were to be solved as rapidly as possible. In fact, analyses of the time for each arithmetic problem during the computation span task indicated that an average of approximately 4 s was spent on each problem, compared with the less than 2 s needed by pilot subjects performing the arithmetic problems without any concurrent memory requirement. Furthermore, the fact that considerably smaller spans are obtained when the task is administered under experimenter-paced conditions of about 2 s per problem (Salthouse, 1988b; Salthouse & Prill, 1987) also suggests that deliberate rehearsal strategies may have been used by at least some of the present subjects.

The possibility that subjects might have varied in the manner in which they performed the task clearly complicates interpretation of the resulting measures as reflections of a working-memory processing resource. The nature of this complication can be elucidated by thinking of working memory as involving the two simultaneous tasks of storage and processing. As with other dual-task situations, therefore, comparisons of performance in the measures of one task are most meaningful only if there is little or no variation in the performance of the other task. Because the lengthy average times taken to perform the arithmetic (processing) task were also associated with considerable between-subject variability, it is likely that subjects differed in their relative emphases on storage versus processing, and thus the CSpan measure may not have reflected the same property of working memory in every individual. One means of investigating this interpretation, currently under way in our laboratory, involves measuring the efficiency of processing and the capacity of storage in both single- and dual-task situations and then examining the relative emphases on processing as opposed to storage in the dual-task, or working-memory, situation.

Still another issue concerning the validity of the CSpan measure relates to the manner in which working memory is most appropriately assessed. That is, it could be argued that working memory implies that the memory is actually used during the performance of a relatively complex cognitive task. From this perspective, therefore, the best measures of working memory may not be those derived from a task specifically designed to measure a particular type of memory capacity, but rather those obtained from assessments within the context of on-going cognitive activities.

In the present study, the slope measures representing the change in accuracy for one-relevant trials across increases in the number of presented premises or folds appear promising as potential candidates for this type of within-context working-memory assessment. The discovery that the values from the verbal reasoning task were significantly correlated with those from the spatial paper-folding task is also encouraging because it suggests that the measures reflect a general, rather than task-specific, construct. Unfortunately, the assessment procedures used in this study resulted in the slope measures having relatively low reliability, and therefore their usefulness in statistical-control analyses is somewhat limited. Reliability could undoubtedly be improved by increasing the number of appropriate observations from each subject, however, and further investigation of within-context working-memory assessment definitely appears warranted.

Conclusion

In conclusion, the results of the present experiment appear reasonably consistent with the processing-resource interpretation of age differences in cognitive functioning. Although this perspective is still uncomfortably vague, it seems to provide a parsimonious account of three results that are currently difficult to explain without the processing-resources construct: (a) increases with age in the magnitude of complexity-related performance differences; (b) high correlations across subjects in the magnitudes of the complexity slopes from the verbal integrative-reasoning task and the spatial paper-folding task; and (c) moderate attenuation of the age effects with statistical control of an index of the working-memory processing resource. It is clear, however, that greater understanding is needed of the nature of working memory, and how it can best be measured, before a consensus can be expected with respect to the contributions that age-related reductions in working-memory processing resources might make to age differences in cognitive performance.

References

Baddeley, A. (1986). *Working memory.* New York: Oxford University Press.

Baddeley, A., Logie, R., & Nimmo-Smith, I. (1985). Components of fluent reading. *Journal of Memory and Language, 24,* 119–131.

Case, R. (1985). *Intellectual development from birth to adulthood.* Orlando, FL: Academic Press.

Case, R., Kurland, M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33,* 386–404.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action.* New York: Houghton Mifflin.

Craik, F. I. M., & Rabinowitz, J. C. (1984). Age differences in the acquisition and use of verbal information: A tutorial review. In H. Buouma & D. G. Bouwhuis (Eds.), *Attention and performance X* (pp. 471–499). Hillsdale, NJ: Erlbaum.

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

Hebb, D. O. (1942). The effect of early and late brain injury upon test scores, and the nature of normal adult intelligence. *Proceedings of the American Philosophical Society, 85,* 275–292.

Horn, J. L. (1982). The theory of fluid and crystallized intelligence in relation to concepts of cognitive psychology and aging in adulthood. In F. I. M. Craik & S. Trehub (Eds.), *Aging and cognitive processes* (pp. 237–278). New York: Plenum.

Salthouse, T. A. (1985). *A theory of cognitive aging.* Amsterdam: North-Holland.

Salthouse, T. A. (1988a). Resource-reduction interpretations of cognitive aging. *Developmental Review, 8,* 238–272.

Salthouse, T. A. (1988b). The role of processing resources in cognitive aging. In M. L. Howe & C. J. Brainerd (Eds.), *Cognitive development in adulthood* (pp. 185–239). New York: Springer-Verlag.

Salthouse, T. A., Kausler, D. H., & Saults, J. S. (1988). Utilization of path-analytic procedures to investigate the role of processing resources in cognitive aging. *Psychology and Aging, 3,* 158–166.

Salthouse, T. S., & Prill, K. (1987). Inferences about age impairments in inferential reasoning. *Psychology and Aging, 2,* 43–51.

Welford, A. (1958). *Ageing and human skill.* London: Oxford University Press.

Received October 26, 1987
Revision received September 1, 1988
Accepted September 1, 1988 ■