Journal of Gerontology
1978, Vol. 33, No. 2, 232-238

The Role of Memory in the Age Decline in Digit-Symbol Substitution Performance[1]

Timothy A. Salthouse, PhD[2]

Adult age differences in the Digit Symbol Test were investigated in three separate experiments. All experiments incorporated a manipulation of the number of digit-symbol pairs in modified versions of the test forms to determine whether the age differences would be reduced with decreased memory requirements. The results of all three experiments indicated that, as the number of digit-symbol pairs was reduced, the age difference was reduced in absolute terms, but not in relative terms. Older subjects decreased their times by greater absolute amounts, but by the same relative amounts when the number of digit-symbol pairs was decreased from nine to one. Independent results supported the interpretation that memory factors play a very small role in contributing to the age decline in digit-symbol performance.

PERFORMANCE on the Digit Symbol Test, one of the subtests in the widely used WAIS intelligence battery, is one of the best psychological correlates of chronological age presently available. Correlations between digit-symbol performance and adult age in large samples of individuals ranging in age from 20 to 90 have consistently been reported to be about -.5 (e.g., Birren & Morrison, 1961; Goldfarb, 1941; Heron & Chown, 1967), and the Digit Symbol Test often exhibits the most dramatic age declines among commonly administered standardized tests (e.g., Hollingsworth & Poffenberger, 1923; Howell, 1955; Hunt, 1949; Whiteman & Jastak, 1957). Despite this impressive status as an age-sensitive performance index, it must be admitted that at the current time there is no adequate explanation of what the Digit Symbol Test measures, nor any compelling hypotheses as to which of the component abilities it requires are responsible for the declines observed with increased age. Some speculations have been offered to the effect that the Digit Symbol Test measures ability to learn (e.g., Thorndike, 1926; Whipple, 1914), translation ability (e.g., Bromley, 1974), or intellectual ability (e.g., Wechsler, 1958), but little or no empirical data have been provided in support of these conjectures.

One interpretation of digit-symbol performance attributes the age decline to a reduction in learning or memory capacity with increased age (i.e., Erber, 1976; Willoughby, 1929). This memory-capacity hypothesis can be tested by varying the number of different digit-symbol pairs presented on the test form and then examining the relationship between age and performance with different numbers of digit-symbol pairs. If most or much of the age difference in digit-symbol performance is caused by a smaller memory capacity in the older individual, then the age difference should be reduced or eliminated as the number of digit-symbol pairs is reduced from nine to six to three to one.

Reducing the number of digit-symbol pairs should not only reduce the number of times the subject has to refer to the code table at the top of the page to retrieve the appropriate code, but it should also reduce the time needed to locate the digit-symbol pair in the table once the search has been initiated. Strictly speaking, then, the manipulation of the number of digit-symbol pairs involves both memory and visual search factors. However, because the time to look from the relevant portion of the response form to the code table can be assumed to be much greater than the time to search the code table once it is fixated, it seems reasonable to postulate that the memory factor is the more important.

[1]This research was conducted while the author was supported by NIA Grant AG-00008 administered by J. Botwinick, Washington Univ., St. Louis. Research funds were provided by The Edwin B. & Frances S. Behrend Memorial Fund for Research on Aging.

[2]Dept. of Psychology, Univ. of Missouri-Columbia, 210 McAlester Hall, Columbia 65201.

The traditional method of determining whether the differences in one factor (e.g., age) vary as a function of the particular levels of another factor (e.g., the number of digit symbol pairs) is to examine the interaction term in an analysis of variance. Representing the two levels of one factor as O and Y, and the two levels of the other factor as A and B, this method of analysis therefore entails the following comparison:

$$O(A) - Y(A) \text{ vs. } O(B) - Y(B) \quad (1)$$

or, equivalently:

$$O(A) - O(B) \text{ vs. } Y(A) - Y(B) \quad (2)$$

It is important to realize that the comparisons involved in this procedure are *absolute differences*. That is, in equation (1) the absolute difference between the Group O and Group Y scores in Condition A is contrasted to the absolute difference between the Group O and Group Y scores in Condition B.

A different perspective to the problem of interpreting age differences across experimental conditions is based on an evaluation of the relationship between average old performance and average young performance in each condition. The rationale behind this approach is that since the absolute levels of performance may vary so widely as to make absolute comparisons of performance quite meaningless, the comparison that should be of greatest interest is the ratio of old performance to young performance in each condition. This leads to comparisons of:

$$O(A)/Y(A) \text{ vs. } O(B)/Y(B) \quad (3)$$

or, equivalently:

$$O(A)/O(B) \text{ vs. } Y(A)/Y(B) \quad (4)$$

An illustration of how these two methods of analyzing age differences can lead to strikingly different patterns of results is available in a comparison of age records for track and field events (i.e., Shepard et al., 1974). The average time for the 100-meter running event was 9.92 sec for record holders between the ages of 20 and 25 and 13.22 sec for record holders between the ages of 60 and 65. The comparable times for the 10,000-meter running event were 1707.5 sec for the 20- to 25-year-olds, and 2236.4 sec for the 60- to 65-year-olds. Substituting these values in equations (1) and (3) reveals that the absolute differences from equation (1) were 3.3 sec and 528.9 sec, but that the relative differences from equation (3) were 1.33 and 1.31. The fact that the performance of the old subjects relative to the performance of the young subjects is the same in the two events is certainly as interesting and important as the fact that the absolute differences in performance increased substantially from the 100-meter event to the 10,000-meter event. Because of the possibility that similar differences in the pattern of results might be obtained in the current experimental tasks, the age differences in the studies reported here will be examined with both absolute and relative comparisons.

EXPERIMENT I

METHOD

Subjects. — Twelve young (age 19 to 28, mean = 23) and 12 old (age 62 to 80, mean = 71) adults served as subjects in a 15-min session following their participation in another unrelated experiment. The young group consisted of 7 males and 5 females, while the old group involved 5 males and 7 females. All subjects reported themselves to be in good health. The young subjects were recruited from the college community and the old subjects from a variety of senior citizen groups and retirement organizations.

Procedure. — All subjects were first administered the standard WAIS Digit Symbol Test and then were tested on eight pages comprising four experimental conditions. Each page had a code table and six rows of 13 double boxes with a digit in the top box and nothing in the bottom box. Two pages contained only one digit-symbol pair in the code table, two pages contained three digit-symbol pairs, two pages contained six digit-symbol pairs, and two pages contained nine digit-symbol pairs. All digit-symbol pairs in the conditions with more digit-symbol pairs included the same pairs that were used in the conditions with fewer pairs. Only the digits contained in the code table on that page were presented in the top boxes on each response form. The digit-symbol pairing in all pages was the same as that used in the WAIS version of the test.

The particular digit-symbol pairs presented in each condition were varied across subjects, and the conditions were presented in a counterbalanced fashion for each subject.

The instructions and procedures throughout all experimental trials were similar to those used in the WAIS Digit Symbol Test, with the exception that there was no time limit and the subjects were required to complete all items on the page in whatever time was necessary.

RESULTS

For ease of comparison, the initial measures of performance on the standard Digit Symbol Test and the modified experimental versions are expressed in seconds/symbol. The mean seconds/symbol on the standard test was 1.26 for the subjects in the young group and 2.33 for the subjects in the old group. A t-test (t (22) = 4.44) indicated that the difference between these values was statistically significant ($p < .01$).

Less than 1% of all responses in the experimental conditions were errors, and thus only correct responses were analyzed. The mean seconds/symbol for the young and old subjects in the four experimental conditions are displayed in the left panel of Fig. 1. An analysis of variance revealed that the age factor ($F(1,22) = 23.03$, MS error = .275), the condition factor ($F(3,66) = 170.28$, MS error = .045) and the age X condition interaction ($F(3,66) = 16.37$, MS error = .045) were all statistically significant ($p < .01$).

The relative method of making age comparisons involves expressing the level of performance in each experimental condition as a percentage of the level of performance in the standard test. (This is simply a special case of equation (4) in which four "A" conditions are all compared to the same "B" condition. The equation (4) version is easier to evaluate statistically than the equation (3) version because separate values are available for each individual in each age group, while equation (3) involves comparing a single average value for each age group.) The means of the relative performance measures for the young and old subjects in the four experimental conditions are displayed in the right panel of Fig. 1. An analysis of variance on these measures indicated that the age ($F(1,22) = 11.24$, MS error = .023) and condition ($F(3,66) = 233.84$, MS error = .010) factors were statistically significant ($p < .01$), but that their interaction was not ($F(3,66) = .56$, MS error = .010).

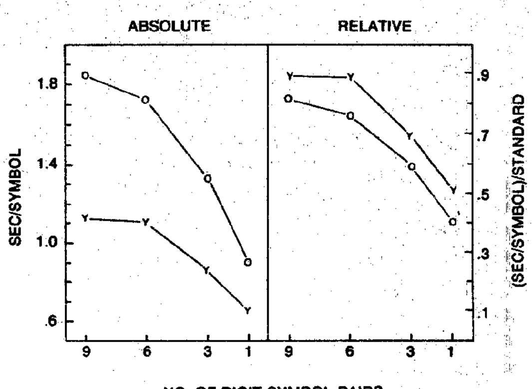

Fig. 1. Absolute (left panel) and relative (right panel) performance for young and old subjects as a function of the number of digit-symbol pairs, Experiment I.

DISCUSSION

The major results of this experiment are subject to two quite different interpretations depending upon which particular dependent measure one chooses to emphasize. The results with the absolute measures of performance indicate that the magnitude of the age difference is reduced as the number of digit-symbol pairs involved in the experimental condition is reduced from nine to one, thus apparently supporting the memory-capacity hypothesis. However, there is no change in the size of the age difference as the number of digit-symbol pairs is reduced with the relative measure, and therefore these results provide no support for the hypothesis.

EXPERIMENT II

The ambiguous status of the memory-capacity hypothesis as a consequence of the different patterns of results with absolute and relative measures of performance makes an independent test of this hypothesis desirable. One such test consists of an examination of the proportion of errors to particular digit-symbol pairs as a function of their frequency of occurrence in the experimental trials. The logic underlying this test is that errors on digit-symbol pairs under speed emphasis instructions can be attributed to two sources: (a) incomplete or inaccurate processing of the digit-symbol pair and (b) failure to retrieve the digit-symbol code from memory. The frequency of the first type of errors can be assumed to remain approximately constant across all digit-symbol pairs, but the frequency of the second type of errors would be expected to be greatest for digit-symbol pairs with the least opportunity to be remembered, i.e., the digit-symbol pairs with low frequencies of occurrence. By manipulating the number of digit-symbol pairs at each of several frequencies of occurrence and examining the proportion of errors in each frequency group, it should be possible to obtain an estimate of the number of digit-symbol pairs remembered by each individual. Assume, for example, that there was one digit-symbol pair with high frequency of occurrence, two digit-symbol pairs with the next highest frequency of occurrence, three digit-symbol pairs with the next frequency, and three digit-symbol pairs with the lowest frequency of occurrence. Now if the proportion of errors for a particular individual was approximately equal for the one, two, and three digit-symbol pairs with high to moderate frequencies of occurrence, but was much greater for the remaining three digit-symbol pairs with low frequencies of occurrence, one could infer that the individual remembered at least six (i.e., one plus two plus three), but less than nine (i.e., one plus two plus three plus three) digit-symbol pairs. This error-proportion test of the memory-capacity hypothesis is examined in the second experiment. In order to produce sufficient numbers of errors, monetary incentives were offered to the subjects to perform at specified rates of speed regardless of the level of accuracy.

Another modification of the procedure from the first experiment involved reducing the writing component in the task by providing symbols in the test spaces below each digit, half of which were the correct symbol and half of which were incorrect, and requiring the subject merely to draw a slash through every incorrect symbol. This modification results in the number of responses remaining constant in all conditions regardless of the number of digit-symbol pairs and thus removes the confounding of number of pairs with number of responses that was present in the first experiment.

METHOD

Subjects. — Thirteen young (age 17 to 23, mean = 20) and 18 old (age 60 to 76, mean = 67) adults participated in a single session of approximately one-half hour. Four of the young subjects and 2 of the old subjects were males. The subjects were drawn from the same population as those in the previous experiment, but none had participated in the earlier experiment. All subjects reported themselves to be in good health.

Procedure. — The procedure was similar to that of the previous experiment with respect to the administration of the standard WAIS Digit Symbol Test, the general construction of the test forms for the experimental conditions, the assignment of digit-symbol pairs to conditions, and the variation of sequence orders across subjects within each age group. The major modifications in procedure concerned the presence of symbols below each digit in the test forms, the partitioning of each test form into 3 trials consisting of two rows each, and the new instructions about varying the emphasis on accuracy.

Unlike the previous experiment, in this experiment a symbol was provided in the box below each digit. The symbols were assigned to the digits such that each digit had a .5 probability of having the appropriate symbol in the box below it and in a manner that would result in exactly half of the digit-symbol pairs in each two-row set being correct. The subject was instructed to draw a slash through all of the symbols that were inappropriate for the digit, according to the digit-symbol code on the top of the form, and to ignore all symbols that were correctly assigned to the digits.

The subjects were offered monetary incentives to perform at different rates of speed in each two-row set on the test form. On half of the pages the subjects performed in the order fast, moderate, and slow, and on the other half they performed in the order slow, moderate, and fast. The speed at which each two-row set was completed was communicated to the subject immediately after its completion. The subjects were instructed to attempt to perform as accurately as was possible consistent with the desired rate of speed. The incentives were based on the subject's time relative to the mean time achieved by the subjects in his age group in a pilot experiment. The slow rate was set equal to the mean speed in the earlier experiment, while the moderate and fast rates were set to be 1.5 and 2.0 times faster than the earlier mean. Despite the incentives being determined by the criterion times that were slower for old subjects than young subjects, the old subjects earned slightly (but not significantly, $p > .05$) less bonus money than the young subjects (i.e., \$0.91 vs. \$1.43, $t(29) = 1.81$).

All subjects received 16 pages of three experimental trials each. The order of presenting the four experimental conditions was counterbalanced for each subject, with different subjects starting with different conditions. The sequence of speed instructions within each condition was also counterbalanced for each subject.

Results

Performance on the standard Digit Symbol Test averaged 1.20 sec/symbol for the 13 young subjects and 1.84 sec/symbol for the 18 old subjects. The t-test ($t(29) = 6.74$) used to evaluate these means revealed that the age difference was statistically significant ($p < .01$).

Because there were substantial differences in both speed and accuracy in each condition, only the trials with 90% or greater accuracy were utilized in the analyses of the speeds. The mean speeds with 10% or fewer errors were determined in each condition for every subject and these mean speeds entered into analyses of variance with both absolute and relative measures of performance. The results with the absolute measures were significant ($p < .01$) effects of age ($F(1,29) = 29.30$, MS error = .095), condition ($F(3,87) = 292.32$, MS error = .039), and of the interaction of age and condition ($F(3,87) = 11.18$, MS error = .039). The condition effect ($F(3,87) = 290.25$, MS error = .016) was also significant ($p < .01$) in the analysis with the relative measures, but the age ($F(1,29) = .02$, MS error = .034) and age X condition ($F(3,87) = .19$, MS error = .016) effects were not. Fig. 2 displays these results.

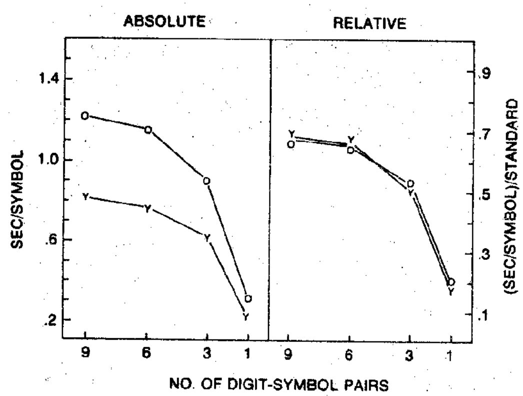

Fig. 2. Absolute (left panel) and relative (right panel) performance for young and old subjects as a function of the number of digit-symbol pairs, Experiment II.

Table 1. Error Proportions as a Function of the Number of Conditions in which the Digit-Symbol Pair Occurred.

	No. of Conditions			
	4	*3*	*2*	*1*
Experiment II				
Young	.139	.140	.177	.182
Old	.083	.084	.163	.145
Experiment III				
Young				
First half	.076	.096	.113	.123
Second half	.084	.088	.131	.126
Old				
First half	.053	.052	.117	.128
Second half	.046	.049	.106	.106

The proportions of errors in the nine-pair condition were computed for each subject for each individual digit-symbol pair, and the pairs were then grouped according to the number of conditions in which they occurred. These proportions were entered into an analysis of variance with age and frequency as factors. The frequency factor proved to be a significant ($p < .01$) source of variance ($F(3,87) = 11.06$, MS error = .009), but the age ($F(1,29) = 4.06$, MS error = .012) and age X frequency ($F(3,87) = .95$, MS error = .009) factors did not. The basis of the frequency effect, as can be seen in the first two rows of Table 1, was for the proportion of errors to be greater for digit-symbol pairs occurring in only one or two conditions (i.e., low-frequency pairs) than for digit-symbol pairs occurring in three or four conditions (i.e., high-frequency pairs). Directional statistical tests on the average error proportion in the two high-frequency conditions versus the average error proportion in the two low-frequency conditions were significant for both the young subjects (i.e., $t(12) = 2.16$, $p < .05$, one-tailed) and the old subjects (i.e., $t(17) = 4.89$, $p < .0005$, one-tailed).

DISCUSSION

The results of this experiment were similar to those of the first experiment in demonstrating that a significant age X condition interaction is evident only in absolute, and not in relative, measures of performance. Unlike the previous experiment, however, in the current experiment an independent test of the memory-capacity hypothesis was available in an examination of the error proportions for digit-symbol pairs of different frequencies of occurrence. Sizeable variations as a function of the frequency of occurrence were apparent in the error proportions (as reflected in the significant frequency effect in the analysis of variance and the pattern of results displayed in Table 1), but there was no evidence of an interaction between age and frequency. This result indicates that the young and old groups of subjects had similar patterns of errors across frequency groupings and, following the argument presented earlier, leads to the inference that the two groups of subjects were operating with functionally equivalent memory capacities.

Because of the importance of the conclusion that young and old individuals perform the Digit Symbol Test with equivalent functional memories, a third experiment was designed to determine whether the error results would be replicated with a new group of subjects. The design of the third experiment was similar to that of the second experiment except that the subjects participated for twice as many experimental trials in order to examine the effects of practice on the variables of interest. The results of this experiment are displayed in Fig. 3 and in the bottom four rows of Table 1. Statistical analyses revealed a pattern of results identical in all important respects to that of Experiment II.

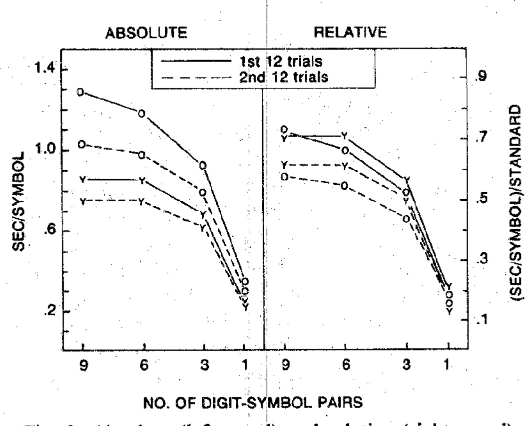

Fig. 3. Absolute (left panel) and relative (right panel) performance for young and old subjects at two stages of practice as a function of the number of digit-symbol pairs, Experiment III.

GENERAL DISCUSSION

There are two major results from the three experiments reported here. The first is the finding, obtained in all three experiments, that as the number of digit-symbol pairs in the condition was decreased from nine to one, the age differences were reduced with the absolute measures, but remained unchanged with the relative measures. The second major result is the observation that while there were systematic variations in error proportions as a function of the frequency of occurrence of the digit-symbol pair, the pattern of errors was the same in the young and old subjects. Both of these results were statistically significant in at least two separate experiments, and thus they provide a firm basis for inference.

The first result is of some general importance since the discovery that age X condition interactions are eliminated with relative measures of performance leads one to speculate whether many of the age X condition interactions previously reported in the literature would also have been eliminated had relative measures been reported rather than absolute measures. The issue as to whether the absolute or the relative measures of performance are the most appropriate for interpreting age differences cannot be resolved here, but it is important to realize that the two measures can, and in the present case do, yield strikingly different patterns of results.

The second result serves to reduce the problem of interpretation created by the discrepant pattern of results with the absolute and relative measures of performance. The inference from the error data, based on the presence of a frequency effect and the absence of an age X frequency effect, is that the young and old subjects performed the task as if they had approximately equivalent memory capacities.

A result reported by Storandt (1976) is also consistent with the hypothesis that memory differences are not responsible for the age differences in digit-symbol performance. Storandt found that old and young subjects require approximately the same proportion of their total digit-symbol time merely to copy symbols already present instead of substituting the symbol appropriate to the presented digit as is required in the traditional form. The copying manipulation is somewhat different from the manipulation employed in the current experiments, but the overall conclusion is the same—reducing the memory requirements in a digit-symbol task does not alter the performance of older adults relative to younger adults.

SUMMARY

Three experiments were conducted to investigate the cause of the age decline in performance on the WAIS Digit Symbol Test. The primary manipulation in all experiments involved varying the number of digit-symbol pairs to be used on a particular test form from the standard number of nine down to six, three, and one. The expectation was that as the number of pairs relevant to the task was decreased, the performance difference between young and old adults would also decrease if memory factors were responsible for the age decline in performance on the standard test. This expectation was confirmed when performance was analyzed in absolute terms, i.e., the actual amount of time required to perform the task, but it was not confirmed when performance was analyzed in relative terms, i.e., the proportion of time required in the experimental conditions compared to the time required in a control condition. Since these two types of analyses yielded contradictory interpretations, an additional test was designed to investigate the issue further. This test was based on the assumption that differences in error proportions across digit-symbol pairs of differing frequencies of occurrence are attributable to memory differences, and thus estimates of functional memory capacity could be determined by examining the number of digit-symbol pairs with relatively low proportions of errors. The results of this test indicated that young and old adults had similar patterns of errors and therefore were presumably performing the digit-symbol task with approximately the same number of digit-symbol pairs in memory. The general conclusion from this study is that the important factors in the age decline in the Digit Symbol Test are perceptual, decisional, or motoric rather than memorial.

REFERENCES

Birren, J. E., & Morrison, D. F. Analysis of the WAIS subtests in relation to age and education. *Journal of Gerontology*, 1961, *16*, 363-369.

Bromley, D. B. *The psychology of human ageing*. Penguin Books, Middlesex, England, 1974.

Erber, J. T. Age differences in learning and memory on a digit-symbol substitution task. *Experimental Aging Research*, 1976, *2*, 45-53.

Goldfarb, W. An investigation of reaction time in older adults and its relationship to certain observed mental test patterns. Teachers College, Columbia Univ., Contributions to Education, No. 831, New York, 1941.

Heron, A., & Chown, S. *Age and function*. Little, Brown, Boston, 1967.

Hollingsworth, H. L., & Poffenberger, A. T. *Applied psychology*. Appleton, New York, 1923.

Howell, R. J. Changes in Wechsler subtest scores with age. *Journal of Consulting Psychology*, 1955, *19*, 47-50.

Hunt, W. L. The relative rates of decline of Wechsler-Bellevue "hold" and "don't hold" tests. *Journal of Consulting Psychology*, 1949, *13*, 440-443.

Shepard, J., Donovan, W., & Mundle, P. *Age records 1974*. Track & Field News, Los Altos, CA, 1974.

Storandt, M. Speed and coding effects in relation to age and ability level. *Developmental Psychology*, 1976, *12*, 177-178.

Thorndike, E. L. *The measurement of intelligence*. Teachers College, Columbia Univ., New York, 1926.

Wechsler, D. *The measurement and appraisal of adult intelligence*. Williams & Wilkins, Baltimore, 1958.

Whipple, G. M. *Manual of mental and physical tests: Part II Complex processes*. Warwick & York, Baltimore, 1914.

Whiteman, M., & Jastak, J. Absolute scaling of tests for different age groupings of a state-wide sample. *Educational & Psychological Measurement*, 1957, *17*, 338-346.

Willoughby, R. R. Incidental learning. *Journal of Educational Psychology*, 1929, *20*, 671-682.