New Tests for Old People

Judith Gonda, Margaret Quayhagen, and K. Warner Schaie

University of Southern California

One of the problems associated with determining the existence and nature of age differences and changes in adult cognitive development across the life-span has been the inappropriateness of existing measures of intelligence for the elderly. Current theory and research suggest tht conventional tests are especially biased against the elderly on two counts. First is the issue of past experience. Since tests were originally designed to predict academic performance in young adults, the abstract materials which comprise most ability tests are quite familiar to many young adults in that their recent school experience has exposed them to regular ability and aptitude testing whereas the elderly are experientially disadvantaged both by being further removed from formal educational experience as well as the fact that even when they were in school the use of tests was not as widely employed as it has been in more recent decades. Consequently, traditional cognitive tasks might be quite familiar to young adults but relatively novel to older adults.

The second issue concerns everyday relevancy or meaningfulness. Having been brought up in the test-taking culture it is conceivable that younger cohorts may be more sympathetic to both the individual and the societal benefits of testing, whereas older cohorts may be culturally alienated to the concept of testing and hence less likely to see reasons for cooperating and becoming sufficiently involved with the task to perform optimally. Thus, traditional cognitive tasks may not be meaningful for older relative to younger cohorts.

Consequently, recent efforts have been made to develop more age-appropriate measures for the elderly (Gonda, 1978; Gonda & Schaie, 1978; Quayhagen, 1978). The present study is a further attempt to evaluate newly-developed measures by comparing the cognitive performance of young, young-old, and old-old individuals on traditional and the most recently developed versions of theoretically more ecologically valid measures (i.e., tasks constructed with familiar rather than novel, abstract materials). Specifically, it was predicted that there would be no differences in spatial and reasoning performance by the young group on traditional versus ecologically valid test forms. Both the young-old and the old-old groups were predicted to perform better on the ecologically valid test forms relative to the more conventional versions, with the performance differential on the traditional versus the familiar test forms expected to be highest in the oldest group.

PROCEDURE

Subjects were solicited from the membership of a local health maintenance organization, and were tested in groups comprised of no more than 20 individuals. Each test session was run by two graduate research assistants, one of which served as an experimenter and the other as a proctor. A total of 264 male and female subjects comprised the three age groups studied. For space, there were 122 old-old subjects (70-75 years), 92 young-old subjects (55-60 years), and 50 young subjects (29-35 years). For reasoning, there were 112 old-old, 95 young-old, and 57 young subjects. Males and females were randomly divided about equally among groups. Four paper-and-pencil tests were used. Tests consisted of two Adult Mental Abilities (AMA) subtests (revised PMA, Thurstone & Thurstone, 1948) for space and reasoning, namely the Figure Rotation (FR) test, and the Letter Series (LS) test. In addition to these conventional tests, two newly constructed ecologically valid space and reasoning analogs were also administered. For space, the Object Rotation Test (ORT) (Quayhagen, 1978) was used and for reasoning, the Word Series (WS) test (Gonda, 1978)

was administered (see Figure 1).

The Object Rotation test was constructed as an alternative and presumably more meaningful form of the PMA Space test. Line drawn familiar objects rather than abstract figures are used as stimuli. Objects were selected for inclusion in the test if they were frequently named, meaningful and easily identifiable household objects. The 20-item ORT was constructed so that the rotated matched or reflected comparison of the objects in each row correspond to the rotated figures of the traditional PMA Space test.

For the Word Series test, the months of the year and days of the week were chosen as verbal stimuli. These words represent familiar overlearned verbal relationships which have smaller range than the 26 letter alphabet, and therefore were assumed to be easier to conceptually manipulate. The repetition pattern for each item of the traditional reasoning test was maintained in the new Word Series test.

The spatial tests were always administered prior to the reasoning tests. Within this fixed order, two orders were used for the administration of the conventional and familiar test forms. In order 1, the AMA subtests preceded the new, ecologically valid tests. In order 2, the reverse was true. Each subject was administered all four subtests under one order, either all in order 1, or all in order 2.

RESULTS

Data for space and reasoning performance were analyzed separately. A 2 x 3 x 2 x 2 (order x age group x sex x test) factorial analysis of variance model was used for both analyses. Repeated measures were made on the test factor.

Space. The spatial data yielded main effects for age, sex and test. Newman-Keuls post-hoc comparisons (Kirk, 1968; Winer, 1971) indicated that all three age groups performed significantly different from one another with highest performance in the young group and poorest in the old-old group. Females scored significantly higher than males, and the new test for space, the ORT, yielded higher performance across age groups. Means and standards deviations for these variables appear in Tables 1, 2, and 3 respectively.

Several significant interactions were also found, however, and thus, some qualifications regarding these main effects are in order. First, an order x age group interaction was found (see Figure 2). Newman-Keuls multiple comparisons made for simple main effects for age group at each level of order substantiated the performance differences between all three age groups at order 2 (ORT administered first) only. At order 1 (AMA administered first), these findings changed slightly since significant differences were found only between young and old-old, and the young and young-old groups. The difference between the old-old and young-old groups diminished at this level of order due primarily to the relatively smaller performance difference in the two orders for the old-old group. Newman-Keuls tests for simple main effects for order at each age group were not significant but both young-old and young groups performed somewhat higher in order 2.

A significant interaction was also found for test x order (see Figure 3). Newman-Keuls comparisons for simple main effects for test revealed significant differences between the two tests at both orders, with the ORT advantage doubling at order 1. Simple main effects for order revealed a significant order effect for traditional PR test. Scores were higher for order 2, i.e., when the traditional test is administered second, suggesting a practice or training effect on the alternate form. There were no significant effects for order on the ORT, although the same pattern occurred, that is, ORT scores were higher when it was presented second also.

Finally, a significant test by age interaction obtained (see Table 4 and Figure 4). Comparisons for simple main effects for test at all age group levels

yielded significant differences with the new ORT yielding higher scores. Simple main effect comparisons for age group yielded a similar pattern of signigicant differences at test 2 (ORT). For the AMA test, however, only differences between young and old-old, and young and young-old groups were significant. Again, this lack of significant difference between old-old and young-old groups appears to be due to the smaller performance differential between the FR and the ORT found in the old-old relative to the other age groups.

Reasoning. Significant main effects for age group were revealed for reasoning performance. As was the case for space, Newman-Keuls comparisons revealed significant differences between all age groups, with the younger groups scoring progressively higher than the older groups. Means and standard deviations can be found in Table 5.

Two significant interactions were also found for reasoning. First, a test x order interaction was found (see Figure 5). Newman-Keuls comparisons for simple main effects for order revealed a significant difference on the new WS test, with better performance for order 1, i.e., when the WS is presented second. The same pattern of results, although nonsignificant, was seen for the traditional LS test, performance was higher when the LS test was presented second, in order 2. There were no significant simple main effects for test at either level of order.

Lastly, the reasoning data also yielded a significant test x age interaction (see Table 6 and Figure 6). Comparisons for simple main effects for age group found all age groups to be significantly different for both tests, with younger groups scoring progressively better. No significant simple main effects for test obtained. However, in contrast to the two older groups, the youngest group scored slightly higher, although not significantly, on the tradition AMA test.

DISCUSSION

The test by age group interactions found in the present study were counter to what had been predicted. For space, simple main effects for test were significant at all three age levels. Significant differences between all three age groups were also found on the ORT, that is, each age group performed significantly better on the new ORT relative to the FR. This finding suggests that the ORT facilitates performance across age groups, and therefore, is probably an easier test than the traditional FR. Thus in this instance, the manipulation of stimulus meaningfulness probably resulted in manipulating the difficulty level as well. Moreover, the fact that the performance of both the young and young-old groups increased on the ORT relatively greater than did that of the old-old group suggests perhaps that this latter group is less susceptible to stimulus manipulation.

For reasoning, the test x age group interaction obtained a different profile to that of space. No significant performance differences on the two test forms were found, although the young group performed better on the traditional form, whereas the two older groups performed about the same on both forms.

The present results are provocative in that attempts to make tests more meaningful apparently is not as simple as substituting familiar items for abstract ones. In the case of space, it appears that increasing the meaningfulness also resulted in an easier test (see Botwinick, 1978). For reasoning, presumably, greater meaningfulness was not attained by the substitution of months and days of the week. Or if it was, it did not affect performance.

The problem remains then, of finding a way in which to make conventional tests more representative of everyday reasoning and spatial behavior while at the same time capturing the exact same level of ability measured by traditional tests. No small feat --- hopefully, the present research will stimulate further research aimed at this challenge.

References

Botwinick, J. Aging and behavior. New York: Springer, 1978.

Gonda, J. Ecologically valid tests and cognitive performance in three age groups. In K.W. Schaie (Chair), Ecologically valid assessment of intellectual functioning in the elderly. Symposium presented at 58th annual meeting of the Western Psychological Association, San Francisco, April 1978.

Gonda, J., & Schaie, K.W. Age-appropriate measures and intellectual competence in three age groups. Paper presented at the 86th annual meeting of the American Psychological Association, Toronto, August, 1978.

Kirk, R.E. Experimental design: Procedures for the behavioral sciences. Belmont, Ca.: Brooks/Cole, 1968.

Quayhagen, M. Test/situational factors in spatial task performance. In K.W. Schaie (Chair), Ecologically valid assessment of intellectual functioning in the elderly. Symposium presented at 58th annual meeting of the Western Psychological Association, San Francisco, April 1978.

Thurstone, L.L., & Thurstone, T.G. SRA primary mental abilities. Chicago: Science Research Associate, 1948.

Winer, B.J. Statistical principles in experimental design. New York: McGraw-Hill, 1971.

Table 1. Means and Standard Deviations for Age Group on Spatial Tests*

Age Group		
Old-Old	Young-Old	Young
18.65	25.01	36.23
(9.62)	(11.10)**	(10.80)

*Total possible score = 54

**Standard Deviations are beneath means

Table 2. Means and Standard Deviations for Sex on Spatial Tests

Sex	
Male	Female
29.67	23.61
(10.67)*	(10.33)

*Standard Deviations are beneath means

Table 3. Means and Standard Deviations for Test Form on Spatial Tests

Test Form	
Figure Rotation	Object Rotation
19.76	27.99
(11.66)*	(13.61)

*Standard Deviations are beneath means

Table 4. Means and Standard Deviations for Age Group x Test Interaction for Space

	Age Group		
Test Form	Old-Old	Young-Old	Young
Figure Rotation	15.66 (8.54)*	19.86 (10.53)	31.21 (11.17)
Object Rotation	21.64 (10.70)	30.16 (11.67)	41.27 (10.42)

*Standard Deviations are beneath means

Table 5. Means and Standard Deviations for Age Group on Reasoning Tests*

Age Group		
Old-Old	Young-Old	Young
8.66 (4.76)	13.40 (5.69)**	19.98 (5.59)

*Total possible score = 30

**Standard Deviations are beneath means

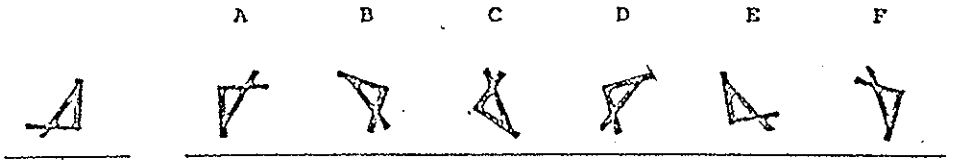
Table 6. Means and Standard Deviations for Age Group x Test Interaction for Reasoning

	Age Group		
Test Form	Old-Old	Young-Old	Young
Letter Series	8.36 (4.85)*	12.84 (5.91)	21.01 (5.90)
Word Series	8.96 (4.67)	13.97 (5.47)	18.93 (5.27)

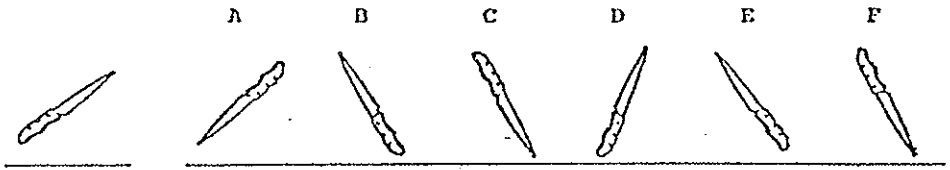
*Standard Deviations are beneath means

Figure 1. Examples of Stimuli

FIGURE ROTATION



OBJECT ROTATION



LETTER SERIES

c d e d e d

ANSWERS

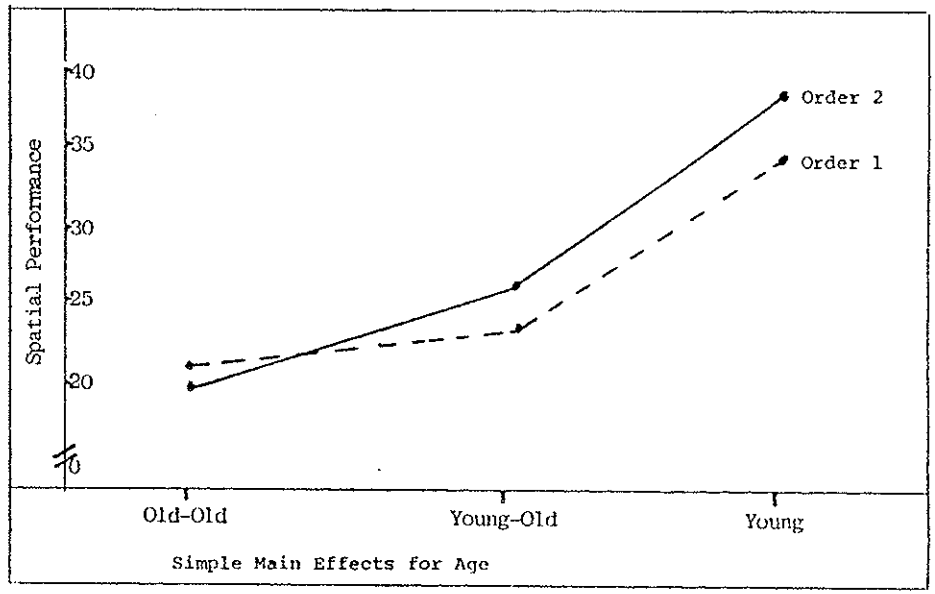
a b c d e

WORD SERIES

ANSWERS

March	January
April	February
March	March
April	April
March	May
April	

Figure 2. Order x Age Group Interaction for Space



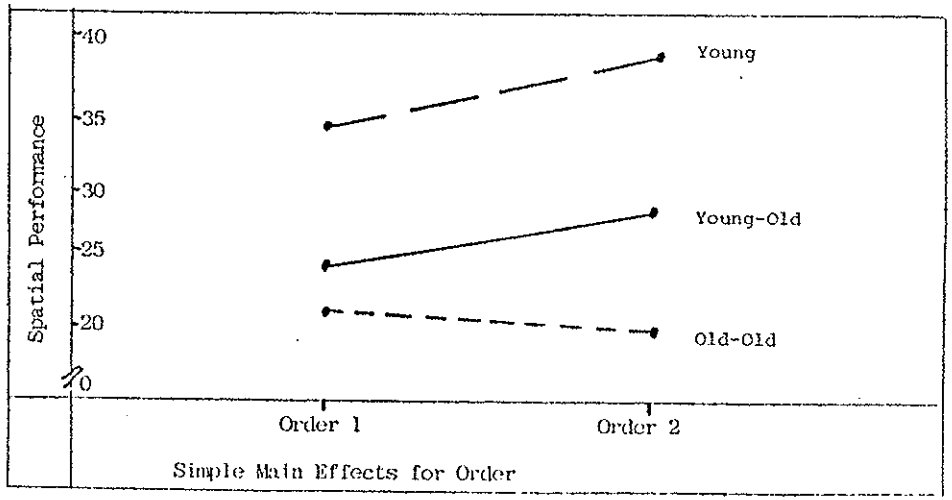
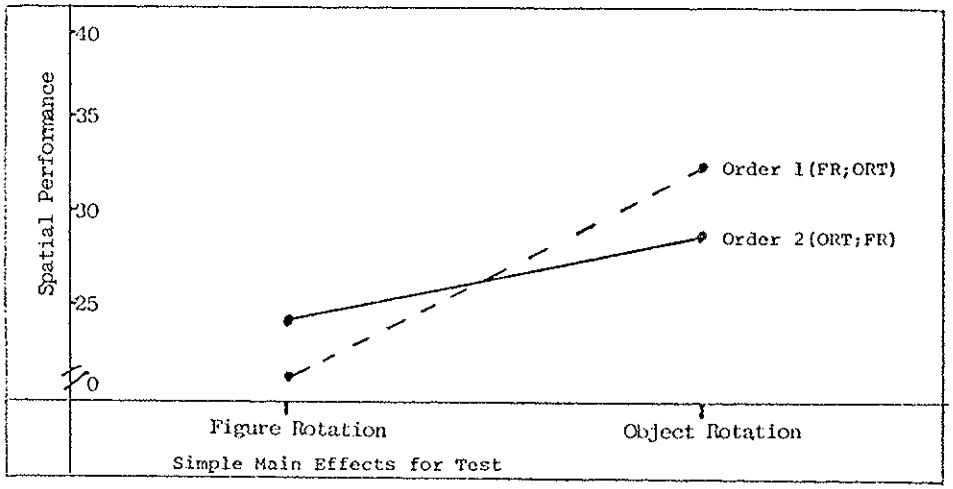


Figure 3. Test x Order Interaction for Space



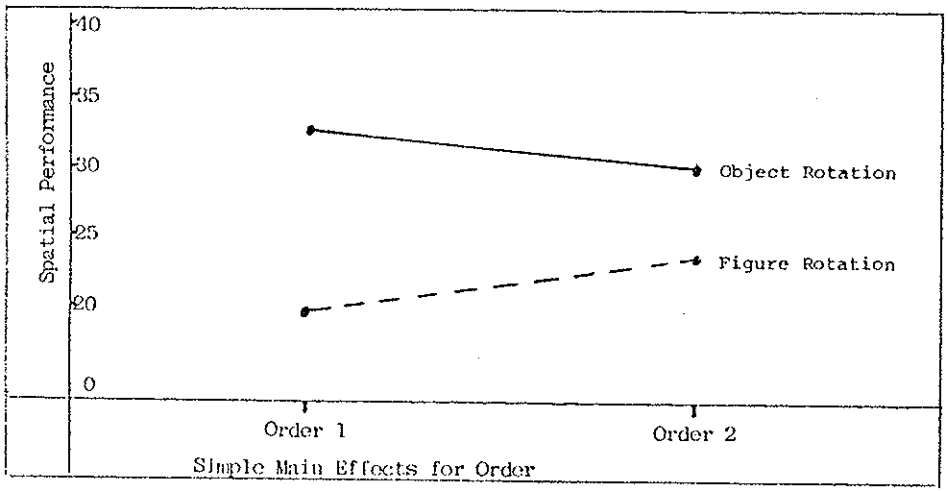
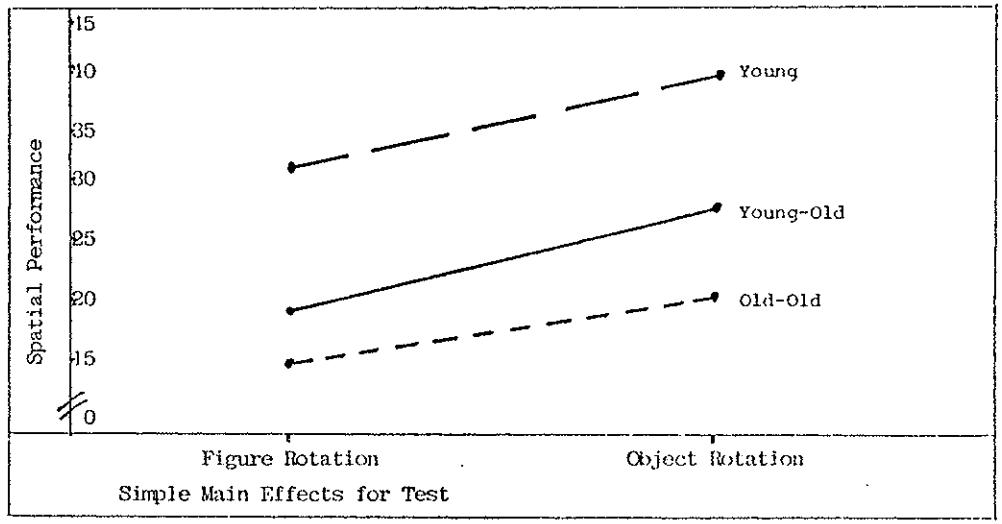


Figure 4. Test x Age Group Interaction for Space



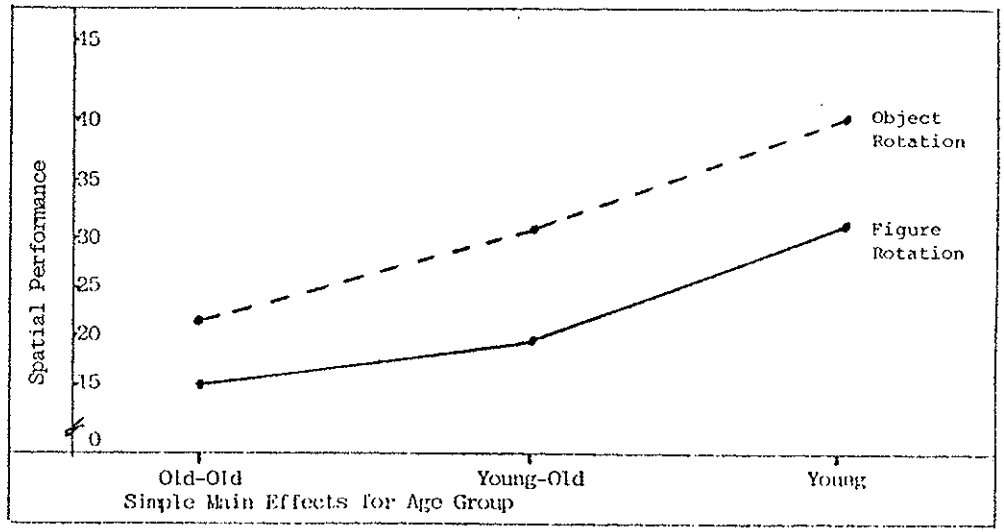
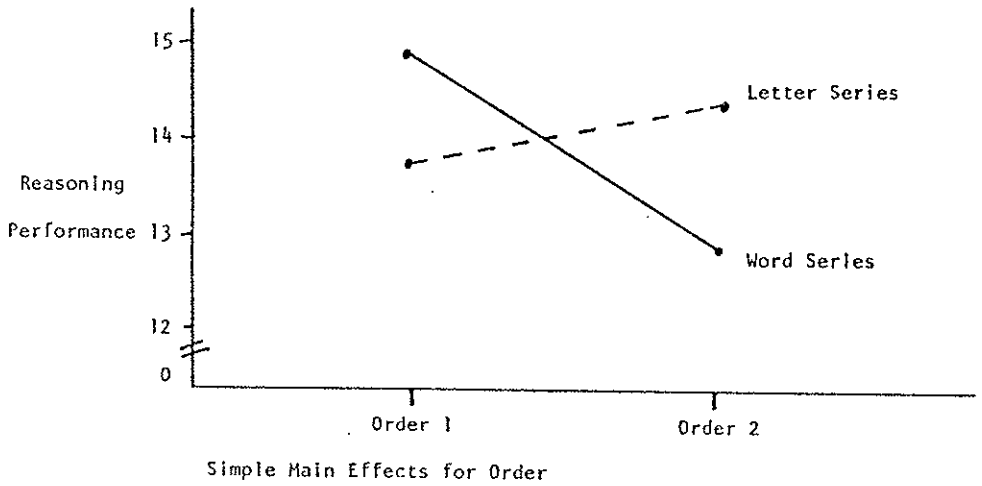
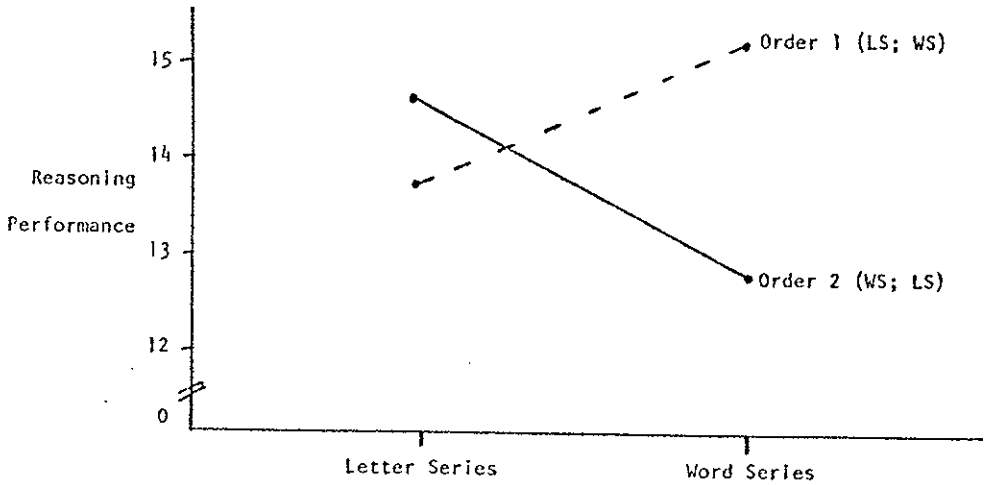


Figure 5. Test x Order Interaction for Reasoning

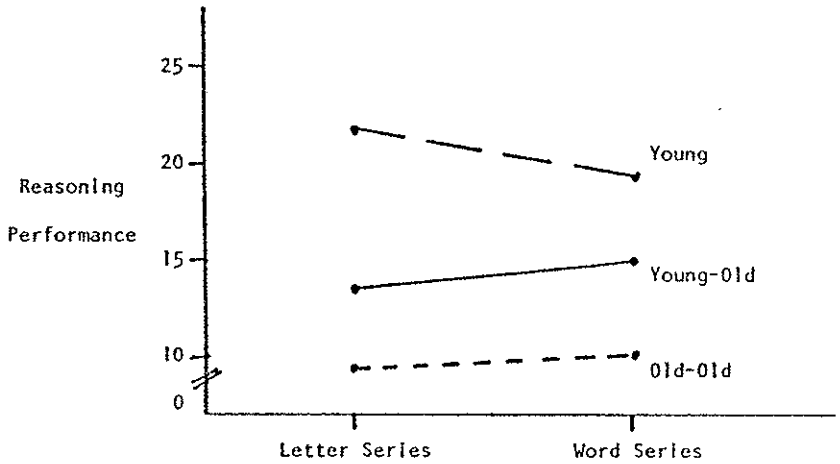


Simple Main Effects for Order

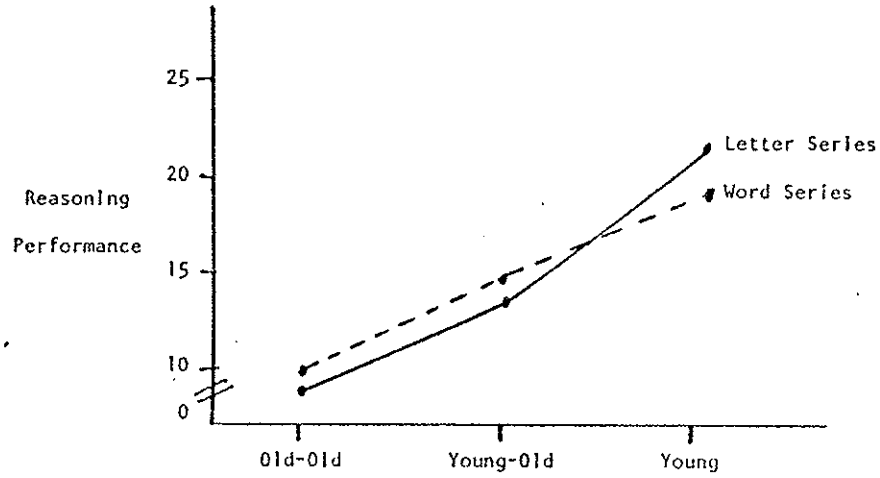


Simple Main Effects for Test

Figure 6. Test x Age Group Interaction for Reasoning



Simple Main Effects for Test



Simple Main Effects for Age Group