

P. M. MOSHER

Sign language training for a
Journal of Geriatric

Received May 3, 1982
Accepted June 28, 1982

AGE-FAIR ASSESSMENT OF PSYCHOMETRIC INTELLIGENCE

Samuel J. Popkin
University of Southern California

K. Warner Schaie
The Pennsylvania State University

Iseli K. Krauss
Syracuse University

This study examined the validity of a parallel form of the Primary Mental Abilities test for use with adults aged 50 to 90 years. Significant intraindividual performance gains on the parallel form were observed on the Verbal and Number subtests, while the effect on Space and Reasoning subtests was less substantial. The development of such age-fair measures of psychometric intelligence and related issues are discussed within the context of current developments in research on aging.

The controversy over stability of psychometric intelligence across the life-span has enjoyed considerable attention in recent years (cf. Baltes & Schaie, 1976; Botwinick, 1977; Horn & Donaldson, 1976). While efforts to discern the effects of normal aging on intellectual development have often emphasized the establishment of appropriate research designs (Schaie, 1965, 1979), there are other important subsidiary issues which require consideration.

One such issue concerns the appropriateness of the actual test instruments employed in the assessment of older adults' intellectual abilities. For example, Marquette (Note 1) indicated that traditional assessment of spatial abilities in older adults may often be more appropriately deemed a test of test-taking ability. Schaie (1978, 1979) and Krauss (Note 2; Krauss, Quayhagen, & Schaie, 1980) have echoed the concern that attention be paid to the test format employed with older adults, lest findings be obscured by the artifact of age- and cohort-related test-taking behaviors.

The present study has as its foundation Schaie's (1979) longi-

This research is an extension of a paper presented at the annual meeting of the Gerontological Society of America, San Diego, California, November, 1980.

tudinal research on the life-span development of psychometric intelligence. The primary assessment tool used in this project has been Thurstone and Thurstone's (1948) Primary Mental Abilities test (PMA), Form AM. While standardized on 11- to 17-year-old students, the PMA has been shown to be a sufficiently difficult test of mental abilities in adults of the fifth to eighth decades (Schaie, Rosenthal, & Perlman, 1953).

An important, albeit informal, observation of the Schaie group has been that adults over 70 years demonstrate considerable difficulty in using the PMA test booklet and accompanying computer-scored answer sheets. Difficulties in actual test-taking behavior seemed to center on the small print size of the materials, unfamiliarity with the use of computer-scored answer sheets, and decreased efficiency in response, given the psychomotor and memory requirement of transferring attention from test booklet to answer sheet.

Given the ample evidence for age-related deficits in visual acuity (see Fozard, Wolf, Bell, McFarland, & Podolsky, 1977 for a review) and psychomotor performance (Welford, 1958; 1977), one may hypothesize that such normal age changes greatly affect the test-taking behavior of older adults. In addition, older adults—many of whom have completed their formal education before the advent of standardized tests—are less likely to be experienced with them. The present study is thus an initial investigation into the development of a more "age-fair" instrument to assess psychometric intelligence. To this end, a parallel form of the PMA was devised, called the Adult Mental Abilities Test (AMA). While test stimuli were identical in both forms, the AMA was designed to eliminate the computer-scored answer sheet, increase the size of printed materials, and further enhance test readability via better spatial blocking of the stimulus items themselves.

The present study attempted to isolate test-taking behavior variables from those cognitive variables which influence measured mental abilities (cf. Krauss et al., 1980). Three specific predictions were entertained. First, it was hypothesized that scores on the PMA subtests which involve basic verbal and numerical skills (i.e., the Verbal-Meaning and Number subtests) would be significantly enhanced through the administration of the AMA. Second, given the observations of the Schaie (1979) group, this performance benefit was hypothesized to be significantly greater for adults of ages 70 to 90 than those of ages 50-69. Third, with respect to performance on subtests requiring abstract reasoning and spatial abilities (i.e., the Space and Reasoning subtests), it was hypothesized that the younger age group might benefit from the AMA format while the older group,

ment of psychometric used in this project has Primary Mental Abilities on 11- to 17-year-old sufficiently difficult test eighth decades (Schaie,

of the Schaie group has considerable difficulty in trying computer-scored ing behavior seemed to be, unfamiliarity with the decreased efficiency in memory requirement of cover sheet.

deficits in visual acuity (Luky, 1977 for a review) (1958; 1977), one may ages greatly affect the tion, older adults—many tion before the advent of rience with them. The into the development of metric intelligence. To revised, called the Adult stimuli were identical in ate the computer-scored materials, and further locking of the stimulus

te test-taking behavior h influence measured er specific predictions at scores on the PMA erical skills (i.e., the ould be significantly MA. Second, given the e performance benefit for adults of ages 70 to t to performance on al abilities (i.e., the and that the younger hile the older group,

experiencing greater difficulty in these novel problem-solving situations (cf. Schaie, 1979), would not obtain a significant performance improvement on the AMA.

METHOD

Experimental Design

A completely crossed, mixed design was employed. The between-subjects factors were Order of test presentations (2), Age (2), and Sex (2). The within-subjects factor was test Form (2).

Subjects

Subjects consisted of 80 volunteer, community-dwelling adults in the Los Angeles area. They were 50–89 years old and assigned to the two Order conditions at random. There were 40 subjects in the *young-old* group (mean age = 62.6 years, $SD = 5.29$) and 40 in the *old-old* group (mean age = 75.3, $SD = 4.48$). There were 10 subjects in each cell of the design. A three-way analysis of variance (ANOVA) revealed no significant differences in education across Order, Sex, and Age. The mean level of education for these study participants was 15.45 years, $SD = 5.08$.

Because of project funding changes during the years in which the data were collected, 50% of subjects were paid \$7 for their participation. This variable was not analyzed in the present study, given the prior research of Gribbin and Schaie (1976) indicating lack of performance differences between paid and unpaid study participants on the PMA variables.

Materials

The PMA and AMA were the tests used in the study. Of the Thurstones' (1948) five PMA subtests, Verbal-Meaning, Space, Reasoning, and Number were administered. The Word-Fluency subtest was not used, since this subtest involves subject generation of written responses on an answer sheet, a task not involving the issues addressed by the present study. Exclusion of this subtest did not affect the order of subtest presentation, since it is the final subtest presented in standard PMA administration. PMA instructions were appropriately altered on the AMA form, indicating that subjects were to respond directly on the test brochure, crossing out the answers they wished to change and re-marking preferred answers.

Procedure

Subjects were tested in groups of four to ten. After a brief introduction, they were asked to record their dates of birth and to indicate the number of years of formal education completed on a form provided in their file folder of test materials. Testing followed immediately thereafter. Subjects in Order 1 were presented the AMA with its appropriate alterations of the 'Thurstones' (1948) PMA instructions. After a 15-minute rest period, the PMA test form was administered. The reverse procedure was applied to Order 2 subjects. Total testing time for each group averaged 2 hours, 15 minutes. The time limits provided by the Thurstones were adhered to: Verbal-Meaning (4 minutes), Space (5 minutes), Reasoning (6 minutes), and Number (6 minutes).

Results

A summary of the major results of the present study is found in Table 1, which lists means and standard deviations of scores across age, test form, and subtest.

Separate analyses of variance were performed for each of the four subtests. Since there were only two levels of the repeated factor in the analysis, it was not necessary to apply the Geisser-Greenhouse correction for heterogeneity of covariance among treatment levels of the repeated factor (cf. Kirk, 1968).

Verbal-Meaning

The main effect for Form was significant, $F(1,72) = 89.71, p < .001$, with the AMA form's eliciting performance superior to the PMA form, as hypothesized. In addition, the main effect for Age was significant, $F(1,72) = 15.31, p < .001$. Subjects in the young-old condition performed significantly better than those in the old-old group. Finally, significant Age \times Form, $F(1,72) = 6.99, p < .01$ and Order \times Form, $F(1,72) = 5.42, p < .025$ interactions were obtained.

Simple main effects were tested to examine the above interaction effects more closely. The AMA-PMA difference for the young-old subjects was significant, $F(1,39) = 25.75, p < .001$ as was the difference for old-old subjects, $F(1,39) = 61.89, p < .001$. Both young-old and old-old groups performed significantly better on the AMA Verbal-Meaning subtest. However, the age-fair version of the PMA particularly benefited the young-old subjects, who appeared to have reached ceiling on this subtest.

to ten. After a brief
ir dates of birth and to
ication completed on a
erials. Testing followed
were presented the AMA
urstones' (1948) PMA
the PMA test form was
ed to Order 2 subjects.
hours, 15 minutes. The
re adhered to: Verbal-
soning (6 minutes), and

sent study is found in
ations of scores across

ned for each of the four
f the repeated factor in
the Geisser-Greenhouse
ong treatment levels of

.72) = 89.71, $p < .001$,
superior to the PMA
an effect for Age was
ects in the young-old
in those in the old-old
(72) = 6.99, $p < .01$ and
ractions were obtained.
the above interaction
nce for the young-old
 $p < .001$ as was the
1.59, $p < .001$. Both
ificantly better on the
fair version of the
who appeared to

TABLE 1 Mean Performance Scores and Standard Deviations for Young-Old and Old-Old Adults for Adult Mental Abilities (AMA) and Primary Mental Abilities (PMA) Tests

Subtest	AMA				PMA			
	Young-old		Old-old		Young-old		Old-old	
	\bar{X}	SD	\bar{X}	SD	\bar{X}	SD	\bar{X}	SD
Verbal	45.2	8.7	39.9	9.9	38.9	8.9	28.5	13.1
Space	23.1	14.5	13.4	9.3	18.2	11.0	10.9	9.2
Reasoning	14.9	6.4	7.7	4.9	14.3	6.3	8.1	5.3
Number	30.9	12.8	28.8	15.6	25.1	11.6	21.7	12.9

Similar tests of simple main effects revealed significant form differences for both Orders 1, $F(1,39) = 22.74$, $p < .001$ and 2, $F(1,39) = 69.79$, $p < .001$. The AMA Verbal-Meaning performance exceeded PMA performance for both orders of presentation, but the difference was slightly greater when the PMA was presented first.

Number

The main effect for Form was significant, $F(1,72) = 64.94$, $p < .001$. As hypothesized, the AMA elicited superior performance to that achieved on the PMA. The prediction that the older group would benefit more from the AMA Number subtest than the young group was not confirmed, but a significant Order \times Form interaction resulted, $F(1,72) = 6.09$, $p < .01$.

The form difference was significant for both orders, $F(1,39) = 21.19$, $p < .001$ for Order 1, and $F(1,39) = 47.94$, $p < .001$ for Order 2. The Order \times Form interaction showed that while the AMA Number subtest resulted in better performance for both Order 1 and 2 groups, the performance gain was somewhat greater when the AMA subtest was presented second.

Space

Significant main effects were observed for Form, $F(1,72) = 30.75$, $p < .001$, Age, $F(1,72) = 14.43$, $p < .001$, and Sex, $F(1,72) = 3.77$, $p < .05$. As predicted, the AMA yielded better performance than the PMA. In addition, the younger group performed better than the older group, and males were found to perform better than females on

this subtest. Significant interactions were obtained for Age X Form, $F(1,72) = 4.11, p < .05$, Order X Form, $F(1,72) = 43.78, p < .001$, Sex X Form, $F(1,72) = 5.99, p < .025$, and Order X Age X Form, $F(1,72) = 5.81, p < .025$.

In the young-old group, the form difference at Order 1 was not significant while significant at Order 2, $F(1,39) = 59.92, p < .001$. In the old-old group, differences at Order 1 were not significant and significant at Order 2, $F(1,19) = 23.58, p < .001$. The Order X Age X Form interaction revealed that, contrary to the original hypothesis of AMA gains by the young-old group only, the AMA Space subtest resulted in better performance for both age groups when it was the second test presented. The benefit to the younger group at Order 2 was, however, nearly twice that of the older group.

Form differences with respect to sex were similarly assessed. The AMA-PMA difference for females was not significant but significant for males, $F(1,39) = 16.63, p < .001$. The AMA form was of benefit to the young-old and old-old males at Order 2.

Reasoning

Contrary to the original prediction, the main effect for Form was not significant. A significant main effect for Age was obtained, however, $F(1,72) = 29.29, p < .001$. The young-old group achieved scores superior to the old-old groups. Significant interactions resulted for Order X Form, $F(1,72) = 38.02, p < .001$, and Sex X Age X Form, $F(1, 72) = 3.20, p < .05$. Simple main effects analyses indicated that form differences were significant for both Order 1, $F(1,39) = 11.23, p < .01$ and Order 2, $F(1,39) = 22.39, p < .001$. Significantly better performance was achieved on whatever Reasoning subtest (AMA or PMA) was presented second.

Tests of simple main effects in the Sex X Age X Form interaction showed the form difference to be non-significant for young-old females, $F(1,19) = .07$ but significant for young-old males, $F(1,19) = 4.47, p < .05$. The difference was not significant for the older females and older males. The younger men did benefit somewhat by the administration of the AMA version, while no other group did so.

DISCUSSION

The results presented above indicate that administration of the AMA version did, as hypothesized, enhance the performance of all subjects on the Verbal-Meaning and Number subtests. Old-old

obtained for Age X Form, $F(1,72) = 43.78, p < .001$, and Order X Age X Form,

ference at Order 1 was not $F(1,39) = 59.92, p < .001$. In $F(1,39)$ were not significant and $F(1,39) = 1.12, p < .001$. The Order X Form interaction was significant, contrary to the original prediction, for both age groups. The benefit to the younger group was greater than that of the older group. The interaction was similarly assessed. The interaction was significant but significant. The AMA form was of benefit at Order 2.

main effect for Form was obtained for Age was obtained, young-old group achieved. Significant interactions were observed for Order 2, $p < .001$, and Sex X Form main effects analyses were significant for both Order 1, $F(1,39) = 22.39, p < .001$. The interaction was observed on whatever Reasoning subtest was presented second.

Sex X Age X Form interaction was non-significant for young-old group. The interaction was not significant. The younger men did not benefit from the AMA version, while

administration of the AMA form reduced the performance of all subtests. Old-old

subjects benefitted more from the AMA on the Verbal-Meaning subtest than did young-old subjects, supporting the original predictions. The differential age benefit was not found, however, for the Number task.

It was further hypothesized that old-old subjects would not benefit by the AMA version on Space and Reasoning tasks, while the younger group would do so. On the Space subtest, both age groups benefitted by the AMA format. The young-old group appeared to benefit from the AMA form more than the older group; however, this difference was observed only when the AMA was presented second. In addition, the male subjects benefitted from the AMA Space form while females did not. A similar pattern was observed on the Reasoning subtest, where the AMA elicited significantly greater overall performance only when presented second, and then only in males in the young-old age group.

Significant test order by form interactions were observed for all subtests administered. On the Verbal-Meaning and Number subtests, there were slight practice effects when the AMA was presented second, although the AMA form clearly enhanced performance at both Orders 1 and 2. Performance gains were observed on the Space subtest only when the AMA was presented second. If practice alone were the sole factor contributing to the observed form difference, it should have operated at both Orders 1 and 2. One may conclude, then, that the AMA form was partly responsible for the significant form difference obtained at Order 2. The only substantial suggestion of uncontaminated practice effects was observed on the Reasoning subtest, where performance gains were obtained regardless of which test form was presented second.

The significant effects of age obtained on the Verbal-Meaning, Space, and Reasoning subtests are difficult to interpret, given the caveats against inferences about age-related decrements in these abilities (Schaie, 1977, 1979). The finding that the young-old group outperformed the old-old group on the AMA form of these subtests suggests that the new parallel form has not compensated for all age-related difficulties which might influence test performance. Which variables may have affected the Verbal-Meaning scores of the older group is unclear; however, the AMA reduced the age difference observed on PMA performance by about 50%. Another finding of interest was the young-old subjects may have reached ceiling on the AMA format. This effect could have suppressed the benefits to be achieved through the parallel form and may raise questions about the suitability of this subtest for use with adults aged 50 to 69 years (Schaie et al., 1953).

The research of Krauss et al. (1980) on the age-appropriate assessment of spatial rotation ability provided evidence that poor memory for actual test stimuli may adversely affect spatial performance, particularly when computer-scored answer sheets are employed. Thus, memory might be one factor which influenced the observed age difference on the Space subtest. Empirical investigations of the influence of psychomotor speed of response and other variables hypothesized to be relevant may be indicated.

In conclusion, it appears that the present initial investigation of the development of age-fair assessment instruments was successful in providing evidence for the utility of such measures' enhancing adult mental performance. It may be feasible to retain within-subjects designs with the refinement of concentrating on each type of mental ability to be assessed, since the AMA appeared to be of a differential benefit as a function of the particular ability assessed. There is a need for additional research which will systematically define the relevant variables accounting for such observed performance differences.

REFERENCE NOTES

1. Marquette, B. W. *Limitations on the generalizability of adult competence across situations*. Paper presented at the annual meeting of the Western Psychological Association, Los Angeles, April, 1976.
2. Krauss, I. K. *Criteria for predictors of competence in the elderly*. Paper presented at the annual meeting of the Western Psychological Association, Los Angeles, April, 1976.

REFERENCES

- Baltes, P. B., & Schaie, K. W. On the plasticity of intelligence in adulthood and old age: Where Horn and Donaldson fail. *American Psychologist*, 1976, 31, 720-725.
- Botwinick, J. Intellectual abilities. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging*. New York: Van Nostrand Reinhold, 1977.
- Fozard, J. L., Wolf, E., Bell, B., McFarland, R., & Podolsky, S. Visual perception and communication. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging*. New York: Van Nostrand Reinhold, 1977.
- Gribbin, K., & Schaie, K. W. Monetary incentive, age, and cognition. *Experimental Aging Research*, 1976, 2, 461-468.
- Horn, J. L., & Donaldson, G. On the myth of intellectual decline in adulthood. *American Psychologist*, 1976, 31, 701-719.

the age-appropriate evidence that poor spatial performance affect spatial performance on answer sheets are not which influenced the test. Empirical evidence for speed of response may be indicated. An initial investigation of the measures' enhancing ability to retain within-contrasting on each type of MA appeared to be of similar ability assessed. It will systematically such observed per-

ity of adult competence meeting of the Western

nce in the elderly. Paper Psychological Association,

intelligence in adulthood *Psychologist*, 1976,

L. K. W. Schaie (Eds.), Van Nostrand Reinhold,

L. Podolsky, S. Visual L. K. W. Schaie (Eds.), Van Nostrand Reinhold,

er, age, and cognition.

Intellectual decline in 19...

- Kirk, R. E. *Experimental design: Procedures for the behavioral sciences*. Monterey, California: Brooks/Cole, 1968.
- Krauss, I. K., Quayhagen, M., & Schaie, K. W. Spatial rotation in the elderly: Performance factors. *Journal of Gerontology*, 1980, 35, 199-206.
- Schaie, K. W. A general model for the study of developmental problems. *Psychological Bulletin*, 1965, 64, 92-107.
- Schaie, K. W. Quasi-experimental research designs in the psychology of aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging*. New York: Van Nostrand Reinhold, 1977.
- Schaie, K. W. External validity in the assessment of intellectual development in adulthood. *Journal of Gerontology*, 1978, 33, 696-701.
- Schaie, K. W. The primary mental abilities in adulthood: An exploration in the development of psychometric intelligence. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life-span development and behavior* (Vol. 2). New York: Academic Press, 1979.
- Schaie, K. W., Rosenthal, F., & Perlman, R. M. Differential deterioration of factorially "pure" mental abilities. *Journal of Gerontology*, 1953, 8, 191-196.
- Thurstone, L. L., & Thurstone, T. G. *Examiner manual for the SRA primary mental abilities test*. Chicago: Science Research Associates, 1948.
- Welford, A. T. *Aging and human skill*. London: Oxford University Press, 1958.
- Welford, A. T. Motor performance. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging*. New York: Van Nostrand Reinhold, 1977.

Received May 22, 1982
Accepted September 10, 1982