Ecologically Valid Tests and Perceived Competence in the Elderly

Judith Gonda, Margaret Quayhagen, and K. Warner Schaie

University of Southern California

In J. D. Sinnott (Chair), Strategies in developing ecologically valid ability measures for mature adults. Symposium presented at the meeting of the American Psychological Association, Montreal, Canada, September 1980.

Ecologically Valid Tests and Perceived Competence in the Elderly

For the past couple of years we have focused our efforts at USC on getting a more accurate assessment of the elderly's intellectual performance by attempting to construct tests that are more meaningful to older individuals. We have undertaken this endeavor with the assumption that part of the age differences observed in studies where the performance of young and old individuals is compared may be attributed to factors that are ability-extraneous but still related to performance. One such factor is the test itself. That is, just because traditional intelligence tests are fairly good predictors of academic performance in younger people it does not necessarily follow that they will also be accurate predictors of everyday competence in older adults.

There are really two issues involved here: 1) whether conventional tests in general are appropriate for use with the elderly, i.e., do they do a good job of accurately assessing the competence of older people? and 2) even if we can devise age-fair intelligence tests, are these tests predictive of everyday competence in the real world of most elderly individuals? In our present study, we think we have taken a preliminary step towards answering these questions.

We have tried to answer these questions by comparing the performance of older people both on traditional and new tests. We then examined the relationship between their performance on these ability tests and their ratings of their own perceived competence in different hypothetical situations in which they might find themselves in the real world. In other words, we essentially asked the question of whether there are differences in the predictive nature of traditional versus new forms of ability tests. And as our criterion measure we used individuals' perception or ratings

of how competently they felt they could deal with various real world situations.

Method

Subjects. Subjects were 403 older men and women who were approximately equally divided into three different age groups for this particular analysis: middle-age (ages 55-64; n=127); young-old (ages 65-74; n=163) and old-old (ages 75-92; n=113). Subjects were solicited from the membership of a health maintenance organization in Southern California. Testing was done in groups ranging in size from 10 to 20 individuals.

Materials. Four paper-and-pencil tests were used for this study. Tests consisted of two Adult Mental Abilities (AMA) subtests which are revised versions of the Thurstone and Thurstone (1948) PMA space and reasoning subtests, namely the Figure Rotation test and the Letter Series test. These subtests have been revised in the sense that we have used larger type size and have modified the format of the tests to be less complex. We have also done away with the machine-scorable answer sheets and instruct subjects to circle the correct answer directly on the test booklet. However, we attempted to keep the difficulty level of each item of our revised versions identical to that of the original PMA subtests. These changes were made because older people may have difficulty with small type and complex formats due to failing eyesight and general unfamiliarity with tests and, consequently, these factors may contribute to the poorer performance of older people.

In addition to these traditional tests, two newly constructed space and reasoning tests that were presumed to be more age-appropriate were administered. For space, the Object Rotation test (Quayhagen, 1978) was

used and for reasoning, the Word Series test (Gonda, 1978). 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 test was constructed so that 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 the days of the week were chosen as verbal stimuli. These words represent familiar, overlearned, verbal relationships which have a smaller range than the 26-letter alphabet and, therefore, were assumed to be easier to manipulate conceptually. The repetition pattern for each item of the traditional reasoning test was maintained in the new Word Series test.

In order to measure individual self-perceptions of competence in various everyday situations the Scheidt (Scheidt & Schaie, 1978) Situational Taxonomy Q-sort was used. This assessment technique consists of 80 cards with a different situation typed on each card. Subjects are asked, for example, such things as how competent they feel filling out a tax form, driving on a freeway, or parallel parking. These situations are classified along eight dimensions which include social/nonsocial, active/passive, common/uncommon, and supportive/depriving. These dimensions are then combined to form 16 classifications (e.g., social-active-common-supportive). Each of these 16 classifications has five situations for a total of 80 situations altogether. The subjects are asked to sort these 80 situations according to the relative degree of competence they feel they have in each situation. The way in which we were then able to relate this assessment

technique to the ability tests was to examine the subjects' ratings of situations within the various dimensions and classifications in relation to their test scores.

Results

A 3x2x2 multivariate analysis of variance with factors of age group (middle-aged, young-old, and old-old), test form (traditional and ecologically valid), and sex was performed on the dependent measures of space and reasoning ability tests. Means and standard deviations are presented in Table 1. Overall significant main effects were found for age group, test form, and sex, along with a signigicant age by test form interaction. An examination of the univariate results revealed that the test form and sex effects and the age group by test form interaction were significant only for the spatial measures. More specifically, performance scores were higher for the ecologically valid form, the Object Rotation test, as compared with the traditional figure rotation test. Men were found to score higher than women, and while the Object Rotation test was easier at all ages, the greatest gain was for the middle-aged group.

Subsequent post-hoc comparisons of mean differences revealed that age differences obtained on all tests between the middle-aged and the old-old, and between the young-old and the old-old. However, a signigicant age difference was found between the middle-aged and the young-old groups only for the Object Rotation test.

Turning to the Situational Q-sort measure, we looked at the mean ratings across age groups on the eight dimensions. Means and standard deviations are presented in Table 2. In general, our findings indicate that social situations (e.g., speaking at a public meeting) are perceived

to demand greater competence than non-social situations, uncommon situations demand greater competence than common situations, and supportive situations demand less competence than depriving circumstances.

In regard to the main question of the relationship between performance on ability tests and perceived competence in everyday situations we will only present some of the general trends we have found among the correlations which were computed between the scores on the four ability tests and the mean ratings on the eight dimensions and some of their permutations. One trend appears to be a positive correlation between ability scores and the active/passive dimension. There also appears to be an inverse correlation between ability scores and the social/nonsocial dimension. The spatial measures tend to correlate inversely with the dimension from social-active-uncommon to nonsocial-passive-common. Reasoning generally correlates positively with the supportive-depriving dimension. This last finding is in agreement with a previous pilot study we conducted in which the Raven's Progressive Matrices was used to measure reasoning ability and, again, reasoning correlated positively with the supportive-depriving dimension in that study (Gonda & Schaie, 1978).

On a more specific level, while at least one of our ability measures predicted some variance in each of the sixteen classifications of the situational taxonomy, the same measure was rarely the best predictor across the different age and sex sub-samples. Consequently, we are left with many different patterns of relationships between performance and perceived competence for the different age and sex sub-groups.

Discussion

Drawing general conclusions from results such as these is not an

easy endeavour. First of all, it is apparent from the findings just reported that our success at constructing a battery of tests to facilitate performance in older adults is only a partial success. While the object rotation test resulted in overall raised performance levels it was in the middle-aged group, rather than in the older groups, that the expected increase was most pronounced. For the reasoning tests, on the other hand, only very slight increases in the new Word Series test over the traditional Letter Series test were observed.

Regarding the suitability of ability tests for predicting everyday competence the present complex pattern of results suggests that perhaps we need to begin to construct tests that are uniquely tailored to measure very specific types of skills in very specific groups of individuals. The ultimate application of tests such as these will be to help in decisions regarding the retention, retraining, or retirement of older workers. While the present study represents some progress toward this goal, clearly we have not yet perfected our methods. Hopefully, the recent call for a more pluralistic approach to ecologically valid inquiry (Scheidt, in press) will generate theorizing and research that will advance both our conceptual and operational thinking in this area.

References

Gonda, J. The word series test. Unpublished test, University of Southern California, 1978.

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

Quayhagen, M. The object rotation test. Unpublished test, University of Southern California, 1978.

Scheidt, R. Ecologically-valid inquiry: Fait accompli? Human Development, in press.

Scheidt, R., & Schaie, K. W. A taxonomy of situations for an elderly population: Generating situational criteria. Journal of Gerontology, 1978, 33, 848-857.

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

Table 1. MEANS AND STANDARD DEVIATIONS FOR THE ABILITY TESTS

		Letter Series		Word Series		Figure Rotation		Object Rotation	
				All Subjects					
Age	f	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
55-64	127	12.54	6.17	13.43	5.57	17.99	11.04	27.95	12.39
65-74	163	9.95	5.71	10.33	5.27	16.00	9.23	23.60	10.96
75+	113	6.39	4.80	7.19	4.68	11.05	10.00	14.84	11.12
				Males Only					
55-64	64	11.70	6.09	13.16	5.34	20.53	10.92	32.66	10.94
65-74	75	9.69	5.29	9.84	5.12	17.75	8.34	25.91	10.92
75+	54	6.17	4.90	6.50	4.66	13.44	10.84	16.39	12.42
				Females Only					
55-64	63	13.38	6.17	13.71	5.82	15.41	10.64	23.17	12.00
65-74	88	10.17	6.06	10.74	5.39	14.51	9.73	21.64	10.67
75+	59	6.59	4.74	7.83	4.66	8.86	8.69	13.42	9.68

Table 2. MEANS AND STANDARD DEVIATIONS FOR THE MAJOR SITUATIONAL DIMENSIONS

	Middle-Aged 55-64		Young-Old 65-74		Old-Old 75+	
	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
All Subjects						
Social (S)	6.26	.37	6.20	.42	6.11	.31
Nonsocial (N)	5.72	.37	5.78	.42	5.86	.31
Active (A)	5.99	.30	5.99	.29	6.05	.27
Passive (P)	5.99	.30	6.00	.29	5.92	.26
Common (C)	5.78	.29	5.68	.42	5.71	.37
Uncommon (U)	6.19	.30	6.29	.43	6.26	.37
Supportive (S)	5.68	.49	5.72	.61	5.81	.49
Depriving (D)	6.28	.49	6.26	.60	6.16	.49
Males Only						
Social (S)	6.26	.37	6.34	.46	6.18	.28
Nonsocial (N)	5.62	.37	5.64	.46	5.79	.29
Active (A)	5.95	.28	6.01	.26	6.03	.31
Passive (P)	6.03	.28	5.98	.29	5.94	.31
Common (C)	5.84	.26	5.77	.54	5.77	.39
Uncommon (U)	6.14	.27	6.21	.57	6.20	.39
Supportive (S)	5.72	.48	5.86	.69	5.91	.51
Depriving (D)	6.26	.48	6.11	.67	6.06	.51
Females Only						
Social (S)	6.16	.35	6.09	.35	6.05	.32
Nonsocial (N)	5.82	.34	5.89	.34	5.93	.32
Active (A)	6.02	.31	5.97	.30	6.07	.21
Passive (P)	5.96	.31	6.01	.31	5.89	.20
Common (C)	5.72	.32	5.60	.25	5.65	.34
Uncommon (U)	6.25	.32	6.36	.25	6.32	.35
Supportive (S)	5.65	.50	5.60	.50	5.72	.46
Depriving (D)	6.31	.50	6.38	.50	6.26	.46