DOCUMENT RESUME

ED 174 683 TM 009 655

AUTHOR Legg, Sue M.; Ware, William B.
TITLE The Role of Fluid, Crystallized, and Creative Abilities in the Prediction of Scores on Essay and Objective Tests.
PUB DATE Apr 79
NOTE 26p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (San Francisco, California, April 9-11, 1979)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Abstraction Levels; *Cognitive Ability; Creative Ability; *Essay Tests; Higher Education; *Intelligence Factors; Intelligence Tests; *Objective Tests; Performance Factors; Predictive Ability (Testing); *Predictor Variables; *Student Ability; Test Results

ABSTRACT

Student and test characteristics were examined by multiple regression analysis and discriminant function analysis to explain why 171 political science undergraduates scored differently on essay versus objective final examinations. Student characteristics included: (1) patterns of creative, crystallized, and fluid abilities as measured by the Cattell Culture Fair Intelligence Test, the McGraw-Hill Basic Skills System Reading Test, and the TORRANCE Tests of Creative Thinking; and (2) previous experience with the subject matter and test format. Test items on the final examination were characterized by the intellectual process required. The essay test was designed to measure creative or fluid (abstract) ability; half of the objective items required fluid ability and half required crystallized (concrete) ability. Results indicated that student characteristics failed to predict total score differences on the final examination, but test characteristics did--students who scored higher on objective tests performed better on abstract items and students who scored higher on essays performed better on concrete objective items. (CP)

The Role of Fluid, Crystallized, and Creative Abilities in the Prediction of Scores on Essay and Objective Tests

by

Sue M. Legg
William B. Ware

Presented at the
National Conference on Measurement
in Education
San Francisco, April, 1979

TITLE: The Role of Fluid, Crystallized, and Creative Abilities in the Prediction of Scores on Essay and Objective Tests

AUTHORS: Sue M. Legg and William B. Ware

INSTITUTION: University of Florida, Gainesville, Florida

This study was designed to assess the contribution of fluid, crystallized and creative abilities to the prediction of success on essay and objective examinations. The characteristics of the student population (N = 171) were assessed. The contribution of the ability measures to the prediction of success on essay and objective tests and on concrete and abstract objective test items was analyzed using multiple regression techniques. A discriminant function was used to analyze the relationship between the ability measures and the classification of students as better on essays, better on objective tests, or the same on both test formats. No differences were found in the pattern of abilities which predicted success on the essay and objective tests or the concrete and abstract items. The discriminant function differentiating the high essay and high objective groups was significant. Students who scored higher on objective tests performed better on abstract items while students who scored higher on essays scored better on concrete objective items.

The Role of Fluid, Crystallized and Creative Abilities in the Prediction of Scores on Essay and Objective Tests

In recent years, psychometricians and cognitive psychologists have begun to conceptualize research questions which span both fields. Concern about the nature of the trait being measured and its relationship to differential cognitive abilities has grown. One of the questions which has arisen from this area of research involves the degree of relationship between types of test items and the mental processes which are required to succeed on the items. However, little empirical research has been done relating differential cognitive abilities to success on essay and objective tests. This specific problem is the focus of this study.

In this study it was anticipated that a combination of student and test attributes would help to explain why many students score differently on essay and objective examinations which were designed to measure the same course content. Two sets of variables were characterized as student attributes. The first set included patterns of fluid, crystallized and creative abilities which predicted test performance (Cattell, 1963). The other set included the previous experience of the students with the subject matter and the test format.

Student attributes alone were not likely to explain the variation in performance on essay and objective tests. Attributes of the tests themselves could influence student performance. Therefore, the test items were classified according to their required intellectual process. The essay item was designed

to measure the ability of the students to synthesize the material and to express any creative insights into the conceptual relationships. One half of the objective test items required the abstract reasoning ability associated with fluid intelligence. The remaining items were designed to measure the ability to comprehend and analyze the subject matter; these skills were more closely related to the crystallized abilities in Cattell's framework. Items associated with fluid ability were labeled abstract, and items which required crystallized abilities were labeled concrete.

This study was designed to evaluate the argument that a broad, unstructured essay measured abstract and creative reasoning abilities while objective tests tended to measure more crystallized abilities. The counter argument that differences in scores on essay and objective tests could be attributed to differences in the abilities required to succeed on abstract and concrete objective test items was also evaluated.

Prior to an examination of the relationships between cognitive abilities and test formats, several measurement problems must be addressed. Early comparison studies between essay and objective examinations yielded substantial correlations between scores on essay and objective tests, but the uncorrelated variance may have been due to either a difference in the function of the tests or to measurement error (Weidemann & Newens, 1933; Vernon, 1959, 1962; Andrews, 1968; Godshalk, Swineford, & Coffman, 1966). Attempts to verify the unique contribution of essay tests by Godshalk et al. (1966), Modu (1972), and Andrews (1968) had contradictory results.

Vernon (1961) stated that it was logical that scores from essay and objective tests over the same content would be correlated. Even though essay questions may have been designed to measure higher level mental processes, evidence indicated that some essays contained substantial factual level information and objective tests may have directly or indirectly measured understanding

and thinking processes. However, both types of tests were imperfect measures.

Measurement errors in objective tests could be traced to a variety of sources: ambiguities due to directions, choice of foils, and relevance to course material and assessment of complex thinking. Vernon cited studies in which objective tests developed by different instructors covering the same material and administered to the same students only correlated about 0.50. Corrections for differences in difficulty of the items and scoring techniques did not account for the low magnitude of these correlations. Vernon concluded that there may be a trade off between the lower reliability of the essay and the validity problems of the objective tests. Moreover, "since the errors which reduce the validities are different it follows that they measure somewhat different aspects of ability" (Vernon, 1961, p. 228). The problem, as Coffman (1971) suggested, is to find valid criteria to document and explain these differences in ability.

The assumption that fundamental differences in cognitive abilities are related to success on essay and objective examinations is implicit in Guilford's (1956) Structure of Intellect model. This model resulted from a factor analytic investigation of a variety of thinking abilities. Memory, cognition, convergent and divergent production, and evaluation factors emerged from the analyses of tests of primary mental abilities.

The Cognition and Production factors aid in the interpretation of the differences in success on various test formats. Guilford identified a type of cognition in which the implications of actions were recognized. This ability was of two types, concrete and abstract foresight. Guilford stated that foresight was an important ability for the political strategist or policy maker. Thus, the extent to which the objective test items measured these cognitive abilities may be important in determining the comparability of the two test formats.

The Production factor in the Structure of Intellect model is particularly relevant to this study. The individual tendency toward convergent or divergent thinking may have a bearing on success on different test formats. Guilford (1956) explained as follows:

> In convergent thinking, there is usually one conclusion or answer that is regarded as unique, and thinking is channelled or controlled in the direction of that answer. In tests of the convergent thinking factors, there is one keyed answer to each item. Multiple choice tests are well adapted to the measurement of these abilities. In divergent thinking, on the other hand, there is much searching or going off in various directions. This is most clearly seen when there is no unique conclusion. (p. 274)

Studies which linked patterns of cognitive abilities and success on different item formats have been encouraged by Carroll (1974) and Messick (1972). Whitely (1976) followed with an investigation of the analogy item in order to further knowledge about abilities and learning. Whitely used the French (1951) kit of primary mental abilities in her analysis. However, Cronbach (1975) advocated the use of broad ability theories in studies of individual differences in learning. Fluid and crystallized abilities are second order factors of Guilford's primary mental abilities. These second order factors provide a broader conceptualization which can provide insight into the patterns of abilities which predict success on different item formats.

An investigation into the relationship of fluid and crystallized abilities and creativity, defined as verbal productive thinking, was conducted by Vernon (1972). In a review of the study, Horn (1976) reported verbal productive thinking did contribute to the prediction of scores on essays and stories beyond the variance due to fluid and crystallized ability.

In a review of the literature on mental abilities, Horn (1976) stated that while verbal productive thinking was largely independent of intelligence, there was doubt about the extent to which it measured real life creativity. He conjectured that when achievement of literary comprehension or critical reading were the dependent variables, a stepwise multiple regression procedure would select crystallized intelligence first, then fluid intelligence followed by a little verbal productive thinking. Deviations from the pattern of regression weights suggested by Horn could indicate that the convergent-divergent thinking factor operates differently on students who score differently on essay and objective tests.

The literature cited supported the hypothesis that different cognitive abilities may be related to success on essay and objective test items. The conceptualization of the patterns of abilities involved has moved from the primary to the more general ability constructs. This study examined the relationship of fluid, crystallized and creative abilities to success on essay and objective tests over the same content.

The Sample

Students enrolled in an introductory political science course were the subjects for this study. There were 171 students enrolled in the course, and 3 students chose not to participate in the study. The remaining students received a bonus for participation to be used in the event their course grade was in question. List wise deletion of cases due to missing data was used.

The course emphasized the study of international relations, and the lectures and examinations stressed conceptual understanding more than factual knowledge. Since the course fulfilled the general education requirement, students were generally sophomores drawn from a variety of departments within the university.

The Instruments

Five instruments were administered: a questionnaire, the Cattell Culture Fair Intelligence Test: Scale Three, the McGraw-Hill Basic Skills Reading Test, the Torrance Tests of Creativity, and a final two hour examination. The final examination included a one hour essay and a fifty item objective test covering the same course content. A two hour essay and objective midterm examination was also administered which gave students experience with the format of the items used on the final examination.

The questionnaire included eight items drawn from suggestions by Coffman (1971). The information obtained from the questionnaire covered the past experience of the students with other political science courses, with the instructor, and with essay and objective test formats.

The final essay test was scored by the instructor on two separate occasions. Global scoring was used. Scoring reliability ($r = .84$) was assessed by correlating the scores on two separate readings of the essay. The criteria developed by the instructor for scoring the essay were:

1. Understanding the dilemma posed by the question
2. Synthesis of diverse material and relevance of examples
3. Discussion and analysis of conceptual issues
4. Originality of perspective

Writing style was a factor in scoring to the degree that good style enhanced the effectiveness of the argument. However, a conscious effort was made to discount grammatical and spelling errors as well as poor penmanship. The essays were shuffled and the names obscured to reduce scoring bias.

The objective test was machine scored. Internal consistency was computed using the KR formula twenty ($r_{xx} = .74$).

Analysis

The analysis of the data was completed in four phases. The preliminary phase involved a description of the relevant characteristics of the population. During the second phase, the contribution of the ability measures to the prediction of success on essay and objective tests was analyzed, and the extent to which the ability measures would predict scores on essay and objective tests was established. Separate multiple regression analyses for each test format were compared using the c matrix of the Biomedical Computer Program (Dixon, 1973). The next question considered the possibility that success on abstract and concrete objective test items would require a different pattern of abilities. Multiple regression was used to analyze the contribution of the ability measures to the prediction of success on the concrete and abstract items. The respective regression weights were compared to determine their similarity.

The premise investigated in the final phase of the study was that differences in abilities could be used to predict relative standing on essay and objective tests. A discriminant function procedure was used to analyze the classification of students as better on essays, better on objective tests or the same on both test formats.

Results

The analyses of the data began with a description of the background characteristics of the sample. The description section was important for two reasons: the sample was not randomly drawn, and previous educational experiences of the students were expected to influence the results of the study.

Background Information

Students were asked to state which test format they preferred and which format was more difficult. An item was also included which asked students

how frequently they tended to write essay and objective examinations. The results for these questions were included in Table 1.

Insert Table 1 about here

Comparison of the Regressions of Essay and Objective Test Scores on the Ability Measures

The multiple correlation coefficients for predicting the objective and essay test scores were significantly different from zero. The multiple correlation coefficient for predicting objective test scores was .46, [F (4,143) = 8.10, $p \leq .05$]. The multiple correlation coefficient for predicting essay test scores by the ability measures was .29, [F (4,143) = 2.75, $p \leq .05$].

The crystallized abilities of comprehension and retention made a statistically significant contribution to the explained variance in the objective and the essay test scores. However, the increases in the R^2 for the measures of fluid ability, originality, and experience, did not approach statistical significance in either equation. The interactions between fluid, crystallized and creative abilities also did not make a significant contribution to the explained variance.

The overall hypothesis of no significant differences in the regression weights between the two equations was supported in the prediction of essay and objective test scores by measures of fluid, crystallized and creative abilities F (4,147) = 2.19, $p > .05$. The standardized regression weights have been reported in Table 2.

Insert Table 2 about here

Predicting Scores on Concrete and Abstract Objective Test Items

The abilities required to succeed on an objective test could be linked to the cognitive level of the objective test item included in the test. The items were categorized as abstract or concrete depending upon the degree of generalization required to respond to the item. The predictor variables (retention, comprehension, and fluid ability) were used to predict scores on the two categories of items. The means and standard deviations for the predictor and the criterion variables have been reported in Table 3.

Insert Table 3 about here

The multiple correlation coefficient for the regression of concrete items on the independent variables was .36 which was significant, F (3,147) = 7.19, $p \leq .05$. The regression of the abstract item scores on the three independent variables produced a multiple correlation coefficient of .45 which reached statistical significance F (3,147) = 12.19, $p \leq .05$.

The Cattell measure of fluid ability did not improve the prediction of scores on the concrete or the abstract items. In fact, the inclusion of fluid ability in the model slightly increased the standard error of prediction for the concrete items.

A summary of the contribution of the predictor variables to the explained variance in the objective test scores was presented in Table 4.

Insert Table 4 about here

The summary of the contribution of the predictor variables to the explained variance in the abstract item scores was included in Table 5.

Insert Table 5 about here

The hypothesis of no significant differences in the comparable regression weights for the concrete and abstract items was supported, F (3,147) = 1.95, $p > .05$. The standardized beta weights for the two equations have been reported in Table 6.

Insert Table 6 about here

Predicting Relative Position on Essay and Objective Tests

The discriminant function analysis was designed to test the degree to which fluid, crystallized and creative abilities would predict the classification of students. The three categories included students whose essay test scores were one z score higher than their objective test scores as one group. The second group included students whose objective test scores were one z score higher than their essay test scores. The third group was composed of all students whose scores on the two forms of the examination were within one z score.

The variables were entered in a direct solution; the combination of variables was expected to predict differences in success on the two test formats. The summary of the statistical tests of significance has been presented in Table 7. The first discriminant function was significant, and its eigen value accounted for 79 percent of the between group variance.

Insert Table 7 about here

The weights for the discriminant functions have been reported in Table 8.

Insert Table 8 about here

The univariate tests for differences among the groups on the independent variables found significant differences between scores on the concrete and abstract items.

The results of the classification analysis indicated that group membership could be predicted for 45.7 percent of the cases. However, much of the error in prediction was attributed to the third group. (See Table 9).

Insert Table 9 about here

Discussion and Conclusions

The overall pattern of the regression of essay and objective test scores on the fluid, crystallized and creative ability scores did not result in statistically significant differences. Crystallized ability scores made the largest contribution to explained variance in both equations. Fluid and creative abilities did not improve the prediction of either essay or objective test scores.

The regression of the scores on abstract and concrete items on measures of fluid and crystallized abilities did not result in significant differences in the overall pattern of the regression weights. Comprehension and retention scores were retained in the model for each equation.

The prediction of relative standing on essay and objective tests resulted in a linear combination of variables which differentiated the abilities of students who scored higher on essay tests from those who scored higher on objective tests. It was the abstract items which differentiated the groups on the discriminant function. Students who scored well on the objective test had higher scores on abstract objective items and on retention. Students who were more successful on the essay also tended to do well on the concrete objective items. This finding contradicts the common wisdom that essays are

the better measures of abstract thinking. In spite of the instructor's stated requirements for the essay, the scores corresponded with the scores on those objective items which indicated mastery of the material rather than abstract reasoning based on conceptual understanding.

Most of the error in prediction occurred in the middle of the distributions of test scores. This lack of homogeneity of error variance could be explained by an analysis of the students' writing ability. It may be that the ability to organize and express ideas clearly was a deciding factor in the scoring of essays in the average essays. Part of the folklore in grading essays is that it is relatively easy to differentiate the excellent and poor papers. The problem in scoring is to separate the average papers from those which are good but not outstanding.

The results of the study indicated that neither measures of fluid, crystallized and creative abilities nor student experiences explained the differences in students' total scores on essay and objective tests. The manipulation of the cognitive requirements of the objective test items on a concrete-abstract dimension did contribute to the explanation of the differences in scores. The logical extension of this study would be to develop the concept of concrete and abstract items in other disciplines and relate scores on these items to scores on essay tests over the same content. Instructors would then be able to more effectively coordinate the testing process with the objectives of the course.

References

Andrews, A. S. Multiple choice and essay tests. Improving College and University Teaching, 1968, 16, 61-66.

Barron, F. Creative person and creative process. New York: Holt, Rinehart and Winston, 1969.

Bloomberg, M. Creativity: Theory and research. New Haven: College and University Press, 1973.

Carroll, J. B. Psychometric tests as cognitive tasks: A new structure of intellect (ETS Technical Report No. 4). Princeton, N.J.: Educational Testing Service, 1974.

Cattell, R. Theory of fluid and crystallized intelligence: A critical experiment. Journal of Educational Psychology, 1963, 54, 1-22.

Coffman, W. Essay examinations. In Thorndike (Ed.), Educational Measurement. Washington, D.C.: American Council on Education, 1971.

Cronbach, L. Beyond the two disciplines of scientific psychology. American Psychologist, 1975, 30, 116-137.

Dixon, W. (Ed.) Biomedical computer programs. Berkeley: University of California Press, 1973.

French, J. The description of aptitude and achievement tests in terms of rotated factors. Psychometric Monographs, 1951.

Godshalk, F., Swineford, F., & Coffman, W. The measurement of writing ability. New York: College Entrance Examination Board, 1966.

Guilford, J. The structure of intellect. Psychological Bulletin, 1956, 53, 4, 267-293.

Horn, J. Human abilities: A review of research and theory in the early 1970's. Annual Review of Psychology, 1976, 27, 437-485.

Messick, S. Beyond structure: In search of functional models of psychological process. Psychometrika, 1972, 37, 4, 357-375.

Modu, C. C. The effectiveness of an essay section in the American history and social studies test (ETS Research Bulletin No. 72-5). Princeton, N.J.: Educational Testing Service, 1972.

Nie, N., Hull, C., Jenkins, J., Steinbrenner, K., & Bent, D. Statistical packages for the social sciences. New York: McGraw-Hill, 1975.

Stenhouse, D. Objective tests, creativity and language games. Australian Journal of Education, 1976, 2, 169-183.

Vernon, P. The assessment of higher intellectual capacities. Oslo: Oslo University Press, 1959.

Vernon, P. The structure of human abilities (2nd Ed.). London: Methuen, 1961.

Vernon, P. The determinents of reading comprehension. Educational and Psychological Measurement, 1962, 22, 269-286.

Vernon, P. The validity of divergent thinking tests. Alberta Journal of Educational Research, 1972, 18, 249-258.

Weidemann, C. C., & Newens, L. F. Does the 'compare and contrast' essay test measure the same mental functions as the true-false test? Journal of General Psychology, 1933, 9, 430-449.

Whitely, S. Solving verbal analogies: Some cognitive components of intelligence test items. Journal of Educational Psychology, 1976, 68, 234-242.

Wittgenstein, L. Philosophical investigations. Oxford: Blackwell, 1953.

Table 1

Summary of Items about Test Format

	N	%
Which test format do you prefer?		
Essay	60	35
Objective	64	38
No preference	45	26
Other	1	1
Have you written essay tests in other courses?		
Usually	53	31
About half of the time	82	48
Seldom	34	20
Never	1	1
Which test format is more difficult for you?		
Essay	48	28
Objective	51	30
Neither	71	42

Table 2

Standardized Beta Weights for

Predicting Essay and Objective Test Scores

Dependent Variable	Independent Variables			
	Retention	Comprehension	Cattell	Originality
Essay	.020	.270	-.055	.095
Objective	.241	.266	.025	.110

Table 3

Summary of the Descriptive

Statistics for the Criterion Variables

	No. of Items	Mean	S.D.
Concrete Items	22	12.59	3.10
Abstract Items	23	12.08	3.09

Table 4

Summary of the Changes in R^2 due to the Addition of Predictor Variables using Concrete Item Scores as the Criterion

	R	R^2	R^2 Change	Simple R
Comprehension	.33	.11	.11	.33
Retention	.35	.12	.02	.27
Cattell	.36	.13	.00	.06
Standard error = .95				

Table 5

Summary of the Changes in R^2 due to the Addition of Predictor Variables using Abstract Item Scores as the Criterion

	R	R^2	R^2 Change	Simple R
Comprehension	.37	.14	.14	.37
Retention	.43	.18	.04	.36
Cattell	.45	.20	.02	.27
Standard error = .91				

Table 6

Standardized Beta Weights for the

Regression of Concrete and Abstract Items on

Measures of Fluid and Crystallized Abilities

	Comprehension	Retention	Cattell
Concrete Items	.274	.161	-.074
Abstract Items	.229	.213	.134

Table 7

Summary of Statistical Tests

Disc. Func.	Eigen Value	Rel. Percent	Canonical Corre-lation	Wilks Λ	X^2	D.F.	Sig.
1	.195	78.96	.40	.795	33.356	12	.001
2	.052	21.04	.22	.951	7.385	5	--

Table 8

Standardized Discriminant Function Coefficients

	Function 1	Function 2
Retention	.1602	-.3586
Comprehension	-.2309	.7200
Originality	-.0571	-.3687
Cattell	-.1162	-.7956
Abstract	1.0384	.0455
Concrete	-.0209	.1326

Table 9

Percentage of Cases Correctly Classified

Group		N	Predicted Group 1	Predicted Group 2	Predicted Group 3
1	N	32	21	4	7
	%		66	12	22
2	N	32	5	19	8
	%		16	59	25
3	N	87	25	33	29
	%		29	38	33