ACT RESEARCH REPORT

No. 30

30

August, 1969

DIFFERENTIAL VALIDITY IN THE ACT TESTS

Nancy S. Cole

Published by Research and Development Division

American College Testing Program

P. O. Box 168, Iowa City, Iowa 52240

DIFFERENTIAL VALIDITY IN THE ACT TESTS

NANCY S. COLE

Summary

The differential validity of subject area tests of academic ability is investigated. Principal components analyses of test scores, high school grades, and college grades in English, math, social studies, and natural sciences show a dominant general ability dimension and a consistent configuration of subject areas on second and third dimensions.

Data from approximately 250 colleges yield correlations of subject area college grades with subject area test scores on the American College Tests and with high school grades. A criterion of differential validity is proposed and calculated for the ACT tests and high school grades in predicting college grades. The moderate differential validity found is interpreted in terms of the first analysis.

Differential Validity in the ACT Tests

Nancy S. Cole[1]

Despite the successes of standardized tests of academic ability, one area has remained a problem. This is the area of differential prediction. The ease with which tests have predicted overall academic success has led to the demand for more specific tests to differentiate ability in various academic areas.

Because of the persuasive content validity of many of these subject area tests, verification of their differential validity has too often been ignored. For example, in his review of the College Entrance Examination Board (CEEB) admissions testing program, Fricke (1965) criticized the Scholastic Aptitude Test (SAT) and the CEEB achievement tests for their lack of differential validity and also noted the relatively little research evidence available.

The purpose of this paper is to investigate the differential validity of one commonly used college admissions test, the American College Test (ACT). Differential validity is of special concern because the relative scores on the four ACT tests in English, mathematics, social studies, and natural sciences are often used for evaluating a student's relative abilities in the four subject areas.

Two important aspects of predictor, x, and criterion, y, behavior are related to differential validity. The first is the degree of the correlations among the variables within the predictor and criterion sets (Guilford, 1956; Thorndike, 1950; Wesman and Bennett, 1951; etc.). When these correlations, r_{x_i,x_j} and r_{y_i,y_j}, are high, the predictors and criteria have little independent variance. Thus when x_i predicts y, x_j also tends to predict it. Similarly, when y_i is predicted by x, then y_j also tends to be predicted by x. The relatedness of the ACT tests and high school grades, the predictors, and of college grades, the criteria, is considered in Study 1.

A second and more direct indicator of differential validity comes from the comparison of the correlations r_{x,y_j} with r_{x,y_k} for the set of predictors (Brogden, 1951; Cronbach, 1960; Horst, 1954; Mollenkopf, 1950; etc.). If x correlates positively with y_j but little or negatively with y_k, then x is a suitable differential predictor for y_j and y_k. In Study 2, these r_{xy} correlations are collected for both ACT tests and high school grades as x and for college grades as y.

Finally, in Study 3, a criterion of differential validity suitable to the differential use of test scores and high school grades is presented. Using data presented in the first two studies, we then evaluated the differential validity of the ACT tests and of high school grades according to the proposed criterion.

---

[1]The author is indebted to James M. Richards, Jr. and Leo A. Munday for their helpful suggestions.

Study 1

As already noted, differential prediction is limited by similarities among the criteria to be predicted and among the predictors. Thus, to evaluate and understand the amount of differential validity in tests of academic ability and in high school grades for predicting college grades differentially, we must first understand the degree of relatedness of the predictor variables and of the criteria.

Data. The American College Testing Program provides research services to ACT-participating colleges. Included in the Standard Research Service analyses are correlations among the college grades in four subject areas (English–E, math–M, social studies–SS, and natural sciences–NS) which the colleges have reported. These correlations were collected for approximately 100 colleges participating in 1968 with a combined N of over 20,000 for each correlation. The average of Fisher's z-transformations weighted by their sample sizes, which were then transformed back to correlations, gave an estimate of the correlations among college grades. The correlations thus found are given in Table 1.

Table 1

Correlations Among College Subject Area Grades

		Col			
		E	*M*	*SS*	*NS*
Col	*E*	–	.43	.52	.50
	M		–	.46	.54
	SS			–	.57
	NS				–

Correlation matrices for self-reported HS grades (in E, M, SS, and NS) and for the four ACT tests (E, M, SS, NS) were available from Holland and Richards (1967) for a large (N=18, 378) representative sample of students taking the ACT test in 1964-65. These correlations are given in Table 2. A correlation matrix for the four ACT tests and the two SAT tests (Verbal–V and Math–M) was reported by Sassenrath and Pugh (1965) for 708 Indiana University students.

Table 2

Correlations* Among ACT Scores and High School Grades

		ACT						HS			
		E	*M*	*SS*	*NS*			*E*	*M*	*SS*	*NS*
ACT	*E*	4.97	.61	.68	.65	HS	*E*	.85	.44	.56	.48
	M	.62	6.38	.61	.62		*M*	.44	.99	.42	.50
	SS	.70	.61	6.28	.75		*SS*	.55	.43	.90	.48
	NS	.66	.61	.74	6.09		*NS*	.46	.46	.49	.92

**Note*—Correlations for men are above the diagonals and for women below. Standard deviations for men are on the diagonals. (From Holland and Richards, 1966, p. 5)

Results. To understand the similarities and differences in the four subject areas (E, M, SS, and NS), we submitted each of the four correlation matrices available (college grades, high school grades, ACT scores, and ACT plus SAT scores) to a principal components analysis. The results were strikingly similar in all four cases. The first root was quite dominant, accounting for from 61% to 74% of the trace. The corresponding loadings were high and positive for all tests or grades (from .74 to .89) indicating many abilities (or one general ability) common to the subject areas.

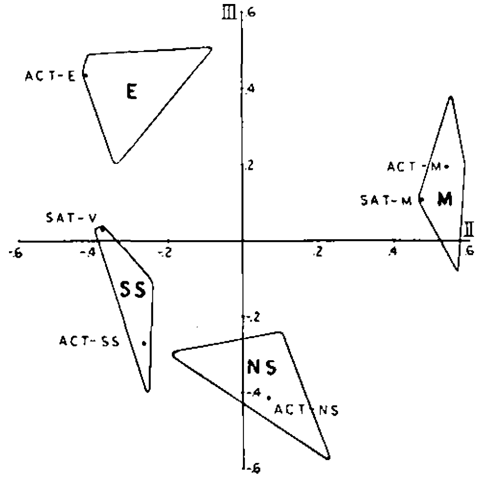

FIG. 1. Subject areas in college grades, high school grades, and tests are plotted on second and third dimensions from principal components analyses.

However, the dimensions on which the remaining small differences existed could be important to the question of differential prediction. The configuration of the subject areas on the second and third dimensions in the principal components analyses are given in Figure 1. All four analyses yielded points within the regions shown. The points for the Sassenrath and Pugh matrix are given to relate the ACT and SAT tests. The reliability of the configuration of subject areas is confirmed by its occurrence in tests, college grades, and high school grades and from data of different sources.

Discussion. From these analyses we know that the tests and the grades they are to predict have sizable similarities. However, consistent though moderately-sized differences in high school and college subject area grades suggest that some differential prediction can occur.

Study 2

To directly assess the differential validity of the ACT tests (and of high school grades for comparison), we obtained the correlations of college grades in the four subject areas (E, M, SS, and NS) with the four ACT tests and high school grades.

Data. Part of the analyses provided colleges by the American College Testing Program research services is the correlations of the ACT scores and self-reported high school grades (routinely reported by the student in the ACT battery) with first semester college grades. For the years 1966 and 1967 the correlations for the approximately 250 colleges with this data were collected. Not all colleges reported grades for all four areas so the numbers in each area differ as follows: E–278 colleges (N=229, 265); M–226 colleges (N=80, 945); SS–261 colleges (N=141, 651); and NS–235 colleges (N=133,702).

Results. The averages of Fisher's z-transformations weighted by the sample size and then transformed back to correlations are given in Table 3. The correlations are lower than would ordinarily be found in an unselected group because of some restriction of range within the colleges, but this should not affect the comparisons made within the table.

Table 3

Correlations of College Grades with ACT Tests and High School Grades

	ACT E	*ACT M*	*ACT SS*	*ACT NS*	*HS E*	*HS M*	*HS SS*	*HS NS*
COL–E	.47	.24	.34	.27	.43	.26	.30	.27
COL–M	.28	.38	.24	.22	.30	.35	.29	.30
COL–SS	.34	.29	.43	.33	.37	.28	.40	.31
COL–NS	.35	.39	.38	.39	.37	.37	.37	.37

Discussion. The amount of differential validity in the ACT scores and in high school grades is indicated by a comparison of a diagonal correlation in each 4x4 matrix with correlations in the column in which it lies. For example, the correlation between ACT–M and COL–M is .38. In the column of ACT–M we find as high a correlation with COL–NS as with COL–M (.39 to .38). Thus, ACT–M gives no differentiation between COL–M and COL–NS although it does differentiate between quantitative and nonquantitative fields.

Table 3 clearly indicates that the ACT tests show as much if not more differential validity than do high school grades. In both cases the amount of differential prediction appears to be moderate.

Study 3

A criterion is needed to make explicit the evaluation of differential prediction. As used here differential prediction refers to the ability of predictors to predict differences in criteria. Thus a reasonable criterion of the differential validity of a pair of predictors (x_i and x_j) for a pair of college grades (y_k and y_h) is the correlation of the difference $d_x = x_i - x_j$ with $d_y = y_k - y_h$. This correlation indicates the degree to which a difference in the predictors is related to a difference in the criteria. In the case of the ACT tests, for example, if a student's English score is higher than his math score, this correlation indicates the degree to which we can expect his college English grade to be higher than his grade in college math.

The correlation of differences, r_{d_x,d_y}, can be expressed algebraically in terms of the variances and covariances of the original four variables:

$$r_{d_x,d_y} = \frac{S_{x_i,y_k} + S_{x_j,y_h} - S_{x_i,y_h} - S_{x_j,y_k}}{\sqrt{(S_{x_i,x_i} + S_{x_j,x_j} - 2S_{x_i,x_j})(S_{y_k,y_k} + S_{y_h,y_h} - 2S_{y_k,y_h})}} \quad (1)$$

for

$$S_{u,v} = \sum_{q=1}^{N} (u_q - \bar{u})(v_q - \bar{v}). \quad (2)$$

The meaning of (1) becomes clear when $S_{x_i,x_i} = S_{x_j,x_j}$ and $S_{y_k,y_k} = S_{y_h,y_h}$, a case which is approximated with the ACT tests, high school grades, and college grades. Then

$$r_{d_x,d_y} = \frac{r_{x_i,y_k} + r_{x_j,y_h} - r_{x_i,y_h} - r_{x_j,y_k}}{\sqrt{2 \quad (1 - r_{x_i,x_j})(1 - r_{y_k,y_h})}} \quad (3)$$

The subscripts in (1) and (3) are written generally to allow for calculation of the differential validity of ACT (SS–NS) for predicting college (E–M), for example. However, in practice, only the corresponding subject areas are usually used.

It should be noted that when y_k and y_h are replications of x_i and x_j (the same tests given at a later time, say), then calling the later tests x_i' and x_j' and the difference $d_{x'}$, $r_{d_x,d_{x'}}$ is a measure of the reliability of difference scores. Letting $r_{x_i,x'_i} = r_{x'_i,x'_j} = r_{x_i,x_j}$, reduces (3) to a familiar form (Gulliksen, 1950),

$$r_{d_x,d_{x'}} = \frac{(r_{x_i,x'_i} + r_{x_j,x'_j})/2 - r_{x_i,x_j}}{(1 - r_{x_i,x_j})} \quad (4)$$

Although our concern is with differential validity, we can appropriately be reminded of the reliability of the differential predictors.

Data. In order to calculate r_{d_x,d_y} for pairs of ACT tests or high school grades and college grades by (3), only the standard deviations of college grades are needed to supplement the data already presented. From a sample of colleges in Study 2 we determined that the standard deviations of college grades within colleges were quite similar to those across colleges. Thus the standard deviations of college grades for students at institutions participating in ACT Research Services in 1965, 1966, and 1967 were used. These standard deviations were .96, 1.18, 1.02, and 1.07 for college E, M, SS, and NS, respectively (Hoyt and Munday, 1968, p. 205).

Results. Although every pair of predictors is a potential differential predictor for a pair of college grades, usually the corresponding pairs (predictors E, M for criteria E, M) were the best. The correlations of differences between predictors with differences between criteria calculated using (3) are reported in Table 4 for ACT scores and in Table 5 for high school grades. The correlations on the diagonals are put in parentheses to identify them as the differential validity of the predictor variables corresponding to the same subject areas as those being predicted. When an off-diagonal validity is greater that the diagonal in its row or column, it is underlined.

Table 4

Correlations of ACT Score Differences with College Grade Differences

		ACT					
		E–M	*E–SS*	*E–NS*	*M–SS*	*M–NS*	*SS–NS*
College	*E–M*	(.33)	.06	.11	–.26	–.22	.05
	E–SS	.19	(.27)	.21	.05	.01	–.05
	E–NS	.29	.18	(.27)	–.12	–.04	.10
	M–SS	–.18	.16	.07	(.30)	.23	–.10
	M–NS	–.07	.11	.14	.16	(.20)	.04
	SS–NS	.12	–.07	.08	–.18	–.05	(.16)

Table 5

Correlations of High School Grade Differences with College Grade Differences

		High School					
		E–M	*E–SS*	*E–NS*	*M–SS*	*M–NS*	*SS–NS*
College	*E–M*	(.19)	.11	.12	–.09	–.07	.03
	E–SS	.06	(.18)	.09	.08	.02	–.07
	E–NS	.16	.14	(.15)	–.04	–.02	.02
	M–SS	–.13	.04	–.05	(.16)	.09	–.08
	M–NS	–.05	.02	.01	.06	(.06)	–.01
	SS–NS	.10	–.02	.07	–.12	–.04	(.09)

Discussion. As suggested in Study 2, the amount of differential validity in ACT tests and high school grades is small. However, it is interesting to note the differences among the combinations. The E, M difference is predicted best, and two other verbal-quantitative differences (E, NS and M, SS) rank high. SS, NS probably ranks low because both the ACT NS test and high school NS are not as mathematical as is college NS. (ACT SS and NS tests are both largely nonquantitative reading tests.) This possibility is confirmed by the fact that SS–M is a slightly better predictor of college SS–NS for both ACT scores and high school grades. A similar situation seems to exist in regard to the low M–NS validity. None of the tests or high school grades correlate better with college NS than any of the others do, and, in particular, ACT NS and high school NS are probably related to college NS mainly through the general ability factor rather than uniquely. In this connection note that E–M predicts college E–NS slightly better than E–NS.

Finally, it should be noted that high school grades are especially poor differential predictors. In addition, the reliability of the difference scores of both test scores and high school grades is low.

Conclusions

This report began by noting the general successes of academic ability testing. From the results presented here, it seems likely that the reasons for those successes are also the reasons for the difficulty in differential prediction. Many areas of academic endeavor involve abilities in common as seen in the first principal component in high school grades, college grades, and test scores. It is these common abilities that are largely measured by tests of academic ability. When differential prediction is desired, the proportion of variation in the predictors and criteria which is unique becomes important. Since the criteria have only a small proportion of variance unique to different subject areas, a test tapping that uniqueness, though suited for differential prediction, would probably have low general validity.

The modest differential validity in the ACT tests and minute amount in high school grades indicate that neither general tests of academic ability nor grades are ideally suited for efficient classification of students into one of several subject area curricula. It seems likely that with special consideration given to differential validity in the construction of the tests, differential prediction could be somewhat improved. However, in tests measuring abilities dominant in college courses, probably only moderate differential validity can be attained. In fact, the goals of general validity and differential validity are apparently in direct conflict in tests of academic ability.

References

Brogden, H. E. Increased efficiency of selection resulting from replacement of a single predictor with several differential predictors. *Educational and Pyschological Measurement* 1951, *11* 173-195.

Cronbach, L. E. *Essentials of Psychological Testing.* New York, New York: Harper and Row, 1960.

Fricke, B. G. Review of the CEEB admissions testing program. In Buros, O. K. (editor), *The Sixth Mental Measurements Yearbook.* Highland Park, N.J.: The Gryphon Press, 1965, 975-988.

Guilford, J. P. *Fundamental Statistics in Psychology and Education.* New York: McGraw-Hill, 1956.

Holland, J. L. and Richards, J. M., Jr. Academic and nonacademic accomplishment in a representative sample of students taking the American College Tests. *College and University, 43,* 1967, 60-71.

Horst, P. A test for the development of a differential prediction battery. *Psychological Monograph,* 1954, No. 380.

Hoyt, D. P. and Munday, L. A. *Your College Freshmen.* Iowa City, Iowa: American College Testing Program, 1968.

Mollenkopf, W. G. Predicted differences and differences between predictions. *Psychometrika,* 1950, *15,* 409-417.

Sassenrath, J. M. and Pugh, R. Relationships among CEEB Scholastic Aptitude Test and American College Test scores and grade point average. *Journal of Educational Measurement, 2,* 2, 1965. 199-205.

Thorndike, R. L. The problem of classification of personnel. *Psychometrika,* 1950, *15,* 215-235.

Wesman, A. G. and Bennett, G. K. Problems of differential prediction. *Educational and Psychological Measurement,* 1951, 11, 265-272.

ACT Research Reports

This report is the thirtieth in a series published by the Research and Development Division of the American College Testing Program. The first 26 research reports have been deposited with the American Documentation Institute, ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington, D.C. 20540. (ADI Document numbers and prices are given below.) Photocopies and 35 mm. microfilms are available at cost from ADI; order by ADI Document number. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress. Beginning with Research Report No. 27 the reports have been deposited with the National Auxiliary Publications Service of the American Society for Information Science (NAPS) c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001. Photocopies and 35 mm. microfilms are available at cost from NAPS. Order by NAPS document number. Advance payment is required. Printed copies are available from the Research and Development Division, American College Testing Program.

The reports are also indexed by the CURRENT CONTENTS, EDUCATION, Institute for Scientific Information, 325 Chestnut Street, Philadelphia, Pennsylvania 19106.

Reports preceded by an asterisk (*) in the list below are available only from ADI or NAPS.

*No. 1 *A Description of American College Freshmen,* by C. Abe, J. L. Holland, S. W. Lutz, & J. M. Richards, Jr.
(ADI Doc. 8554; photo, $8.75; microfilm, $3.00)

*No. 2 *Academic and Nonacademic Accomplishment: Correlated or Uncorrelated?* by J. L. Holland, & J. M. Richards, Jr.
(ADI Doc. 8555; photo, $3.75; microfilm, $2.00)

*No. 3 *A Description of College Freshmen: I. Students with Different Vocational Choices,* by C. Abe, & J. L. Holland
(ADI Doc. 8556; photo, $7.50; microfilm, $2.75)

*No. 4 *A Description of College Freshmen: II. Students with Different Vocational Choices,* by C. Abe, & J. L. Holland
(ADI Doc. 8557; photo $7.50; microfilm, $2.75)

*No. 5 *A Description of Junior Colleges,* by J. M. Richards, Jr., L. M. Rand, & L. P. Rand
(ADI Doc. 8558; photo, $3.75; microfilm, $2.00)

*No. 6 *Comparative Predictive Validities of the American College Tests and Two Other Scholastic Aptitude Tests,* by L. Munday
(ADI Doc. 8559; photo, $2.50; microfilm, $1.75)

No. 7 *The Relationship Between College Grades and Adult Achievement: A Review of the Literature,* by D. P. Hoyt
(ADI Doc. 8632; photo, $7.50; microfilm, $2.75)

*No. 8 *A Factor Analysis of Student "Explanations" of Their Choice of a College,* by J. M. Richards, Jr., & J. L. Holland
(ADI Doc. 8633; photo, $3.75; microfilm, $2.00)

*No. 9 *Regional Differences in Junior Colleges,* by J. M. Richards, Jr., L. P. Rand, & L. M. Rand
(ADI Doc. 8743; photo, $2.50; microfilm, $1.75)

No. 10 *Academic Description and Prediction in Junior Colleges,* by D. P. Hoyt, & L. Munday
(ADI Doc. 8856; photo, $3.75; microfilm, $2.00)

No. 11 *The Assessment of Student Accomplishment in College,* by J. M. Richards, Jr., J. L. Holland, & S. W. Lutz
(ADI Doc. 8955; photo, $3.75; microfilm, $2.00)

No. 12 *Academic and Nonacademic Accomplishment in a Representative Sample taken from a Population of 612,000,* by J. L. Holland, & J. M. Richards, Jr.
(ADI Doc. 8992; photo, $3.75, microfilm, $2.00)

No. 13 *The Prediction of Student Accomplishment in College,* by J. M. Richards, Jr., J. L. Holland, & S. W. Lutz
(ADI Doc. 9020; photo, $5.00; microfilm, $2.25)

No. 14 *Changes in Self-Ratings and Life Goals Among Students at Colleges with Different Characteristics,* by R. W. Skager, J. L. Holland, & L. A. Braskamp
(ADI Doc. 9069; photo, $3.75; microfilm, $2.00)

No. 15 *Can Computers Write College Admissions Tests?* by J. M. Richards, Jr.
(ADI Doc. 9174; photo, $2.50; microfilm, $1.75)

No. 16 *Changes in Self-Ratings and Life Goals as Related to Student Accomplishment in College,* by R. W. Skager, & L. A. Braskamp
(ADI Doc. 9214; photo, $2.50; microfilm, $1.75)

No. 17 *Family Income and the Characteristics of College-Bound Students,* by L. L. Baird
(ADI Doc. 9378; photo, $3.75; microfilm, $2.00)

*No. 18 *Predicting a Student's Vocational Choice,* by J. L. Holland, & S. W. Lutz
(ADI Doc. 9433; photo, $2.50; microfilm, $1.75)

No. 19 *The Educational Goals of College-Bound Youth,* by L. L. Baird
(ADI Doc. 9472; photo, $5.00; microfilm, $2.25)

No. 20 *Who Goes Where to Junior College?* by J. M. Richards, Jr., & L. A. Braskamp
(ADI Doc. 9571; photo, $3.75; microfilm, $2.00)

*No. 21 *Predicting Student Accomplishment in College from the ACT Assessment,* by J. M. Richards, Jr., & S. W. Lutz
(ADI Doc. 9594; photo, $6.25; microfilm, $2.50)

*No. 22 *The Undecided Student: How Different is He?* by L. L. Baird
(ADI Doc. 9812; photo, $3.75; microfilm, $2.00)

No. 23 *The Effects of Selecting College Students by Various Kinds of High School Achievement,* by L. L. Baird & J. M. Richards, Jr.
(ADI Doc. 9955; photo, $3.75; microfilm, $2.00)

No. 24 *Do They Do What They Say They Will Do?* by S. W. Lutz
(ADI Doc. 9988; photo, $5.00; microfilm, $2.25)

No. 25 *Changes in the Vocational Plans of College Students: Orderly or Random?* by J. L. Holland & D. R. Whitney
(ADI Doc. 10051; photo, $3.75; microfilm, $2.00)

No. 26 *The Flow of High School Students to Schools, Colleges, and Jobs,* by L. L. Baird, & J. L. Holland
(ADI Doc. 10053; photo, $6.25; microfilm, $2.50)

No. 27 *Forecasting Academic Success in Specific Colleges,* by D. P. Hoyt
(NAPS No. 00254; photo, $3.00; microfilm, $1.00)

No. 28 *A Description of Graduates of Two-Year Colleges,* by L. L. Baird, J. M. Richards, Jr., & L. R. Shevel
(NAPS No. 00306; photo, $3.00; microfilm, $1.00)

No. 29 *An Empirical Occupational Classification Derived From A Theory of Personality and Intended for Practice and Research,* by J. L. Holland, D. R. Whitney, N. S. Cole, & J. M. Richards, Jr.
(NAPS No. 00505; photo, $3.00; microfilm, $1.00)