NPS 54-81-001

NAVAL POSTGRADUATE SCHOOL

Monterey, California

The Prediction of College Achievement from the Scholastic Aptitude Test and the High School Record

by

R. A. Weitzman

February, 1981

Prepared for:
Office of Naval Research, Arlington, VA 22217

NAVAL POSTGRADUATE SCHOOL

Monterey, California

Rear Admiral J. J. Ekelund
Superintendent

D. A. Schrady
Acting Provost

This report received no direct research support.

R. A. Weitzman, Associate Professor

Reviewed by:

Carl R. Jones, Chairman
Dept. of Administrative Sciences

Released by:

William M. Tolles, Dean of Research

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPS 54-81-001	2. GOVT ACCESSION NO. AD-A133 998	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) The Prediction of College Achievement from the Scholastic Aptitude Test and the High School Record		5. TYPE OF REPORT & PERIOD COVERED Technical, February 1981 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) R. A. Weitzman		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, CA 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE February 1981 13. NUMBER OF PAGES 22
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Predictive Validity; Selection; Incremental Validity; Aptitude Tests; Scholastic Aptitude Test; Part Correlation; College Grading Standards

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Attempting to take a nonadversarial approach to the dispute of Slack and Porter with Jackson over the predictive validities of the Scholastic Aptitude Test and the high school record, this report examines both the effect of the selection process on these validities and what these validities might be if all colleges were to use a common standard of academic achievement. The results indicate that (a) for this criterion the Scholastic Aptitude Test may have a predictive validity of .62 and an

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

incremental validity over the high school record of .17 and (b) for the usual criterion the selection process may have a profound effect on the relative predictive validities of the Scholastic Aptitude Test and the high school record. The conclusion thus reached is that the Scholastic Aptitude Test is, as intended by its developers, a valid instrument of selection.

The Prediction of College Achievement from the Scholastic Aptitude Test and the High School Record

Modern societies invest in higher education largely to promote technologico-cultural advancement and upward socio-economic mobility. The controversy over the use of aptitude tests in the college-admissions process reflects different emphases on these two functions of higher education. The defenders of this use tend more to value the first, the challengers the second. A theory of mental tests exists (e.g., Lord and Novick, 1968), and within this theory the differences between the two sides are arguable in objective terms. Outside the theory, however, the differences consist only of rival values. The public and publicized portion of the controversy has thus tended to be more adversarial than objective or dispassionate.

The two sides collide particularly on the issue of merit versus need as the basis of financial aid for higher education. On this issue, the socio-economic side has increasingly prevailed in the United States. On the issue of college admissions, however, the two sides are currently at a stand-off, though the momentum is against the technologico-cultural side. Exemplifying this tendency are the article by Slack and Porter (1980) and its reply by Jackson (1980), which appeared in the Harvard Educational Review.

The intent here is to focus on one point at issue in these two articles without taking the adversarial position of either. This

issue, the relative validities of the high school record (HSR) and the verbal portion of the Scholastic Aptitude Test (SAT) for the prediction of first-year college achievement (FCA), is resolvable in the context of mental-test theory.

The Opposing Positions

Following Nairn (1980), Slack and Porter (1980) cited the difference in predictive validity of the SAT and the HSR as a compelling reason to favor the use of one rather than the other, or their combination, to aid in the selection of college applicants. The representative predictive validities cited are .37 for the SAT and .52 for the HSR (on a scale of 0, for random prediction, to 1, for perfect prediction). The correspondingly representative predictive validity of the most predictive weighted average of these two predictors is .58. The incremental validity of the SAT over the HSR is thus only .06. These data, in view of the apparent tendency of the SAT to favor upper over lower socio-economic groups more than the HSR does (Goldman and Widawski, 1976), led Slack and Porter to endorse the use of high school grades or other pre-college measures of achievement, rather than aptitude, to predict success in college.

In reply, Jackson (1980) pointed out that the utility of a predictor depends not only on its predictive validity but also on the improvement in criterion performance expected from its use in selection. According to Jackson, the increase from 14.6 to 18.5 in the function of predictive validity that Slack and Porter used to

measure the incremental utility of the HSR plus the SAT over the HSR alone amounts to 27 percent ([18.5 - 14.6]/14.6), not the mere 4 percent (18.5 - 14.6) claimed by Slack and Porter. Because the grading standard varies from high school to high school, Jackson thus endorsed the continued use of the SAT, in combination with high school grades, to predict college success.

Far from objective, this controversy is adversarial to the point of manifest distrust. The opposing sides center in the Ralph Nader and the Educational Testing Service organizations. The data used by Slack and Porter come from a report by Ford and Campos (1977a) prepared for the ETS parent body, the College Entrance Examination Board. Siding with Ralph Nader, the publisher of the Nairn study, Slack and Porter reworked these data, evidently in the belief that Ford and Campos had biased their presentation of them to favor ETS. The .37, .52, and .58 predictive validities presented as representative by Slack and Porter correspond to the over-years median values of .40, .50, and .58 reported by Ford and Campos. Educational Testing Service (1980, Table 1) itself has presented what, except for a lack of good faith, might be generally acceptable as authoritative median figures: .41 for the SAT, .52 for the HSR, and .58 for their combination. The incremental validity of the SAT over the HSR based on these figures is still only .06. Thus by no means self-serving, these figures will provide the basis of all subsequent calculations presented in this report. A choice of figures is necessary even though in this adversarial situation no single choice might seem fair to both

positions.

Resolution of the Dispute

A college-admissions officer reviewing these two positions might well have trouble deciding in favor of one or the other. Even if the incremental validity of the SAT appears to be less than moderate, doubts about the fairness of varying high school grading standards might compel its continued use. Counter doubts arising from the apparent bias of the SAT in favor of upper over lower socio-economic groups might, at the same time, drive the admissions officer in the opposite direction. The vacillation resulting from the opposing forces could certainly benefit from a resolution favoring one side or the other.

The resolution cannot depend on the unreliability of college grades, due to grade inflation, or the restriction in range of SAT scores and high school grades, due to selection, because these conditions affect both the SAT and the HSR predictive validities more or less equally. The resolution offered here depends, rather, on the distinction between a simple correlation and a part correlation. A simple correlation is a measure of the tendency of measurements on two variables to go up or down together; a part correlation is a measure of the tendency of measurements on one variable to go up or down with the parts of measurements on another variable that vary within subpopulations defined by a third variable. The third variable here is college, and the part measurements are the observed within-college first-year grade-point

averages. Complementary to these are unobserved among-college measurements of first-year academic achievement that do not vary for students in the same college but that generally do vary for students in different colleges on a common-standard scale showing college-to-college grade correspondences. This scale might show, for example, that an A in one college corresponds to a B in another. A student's achievement measurement on this scale is equal to the sum of its within-college and among-college parts. The predictive validities of .41 for the SAT and .52 for the HSR are not simple correlations but part correlations—representative correlations between the predictors and FCA (first-year college achievement) measured by a within-college grade-point average. The corresponding simple correlations may be—and indeed estimation later in this section will show them to be—markedly different.

Dependence of the Observed Predictive Validities on the Relative Importance of the Predictors in Selection

Different from their simple-correlation counterparts, the predictive validities of the SAT and the HSR for the within-college measure of FCA depend on the roles of the two predictors in the selection process. Figure 1 illustrates this dependence. The center vertical line represents the scale of SAT scores, and the I, II, III, and IV identify different colleges. The left side describes the part correlations of high school grades (a, b, c, d) and SAT scores with college grades (A, B, C, D) when the SAT is the sole academic instrument of selection; the right side describes the same correlations when high school grades are the sole academic

instrument of selection. On the left side, which more accurately portrays current practice, the relationship between high school and college grades is perfect--a corresponds to A, b to B, c to C, and d to D in each college--while the relationship between SAT scores and college grades is random--350 corresponds equally to A, B, C, and D, as does each of the other SAT scores shown; on the right side, by contrast, the relationship between high school and college grades is random while the relationship between SAT scores and college grades is perfect. On the left, the SAT sorts students into colleges where they can get the same grades that they got in high school; on the right, the HSR sorts students into colleges where their grades can reflect their SAT scores. Because almost every college applicant is acceptable to at least one college, selection is essentially a process of sorting students into colleges. The difference between the representative predictive validities of the SAT (.41) and the HSR (.52) thus may largely reflect the greater weight that colleges have tended to place on SAT scores relative to high school grade-point averages in the selection process. If colleges were to reverse the weights, the difference would be correspondingly in the opposite direction.

Estimation of Predictive Validities for College Achievement Measured on a Common College Scale

The predictive validities illustrated in Figure 1 are extreme cases of the within-college part correlations of FCA with the SAT (.41) and the HSR (.52). The corresponding simple correlations (r_{SF} and r_{HF}) do not depend on the roles of the two

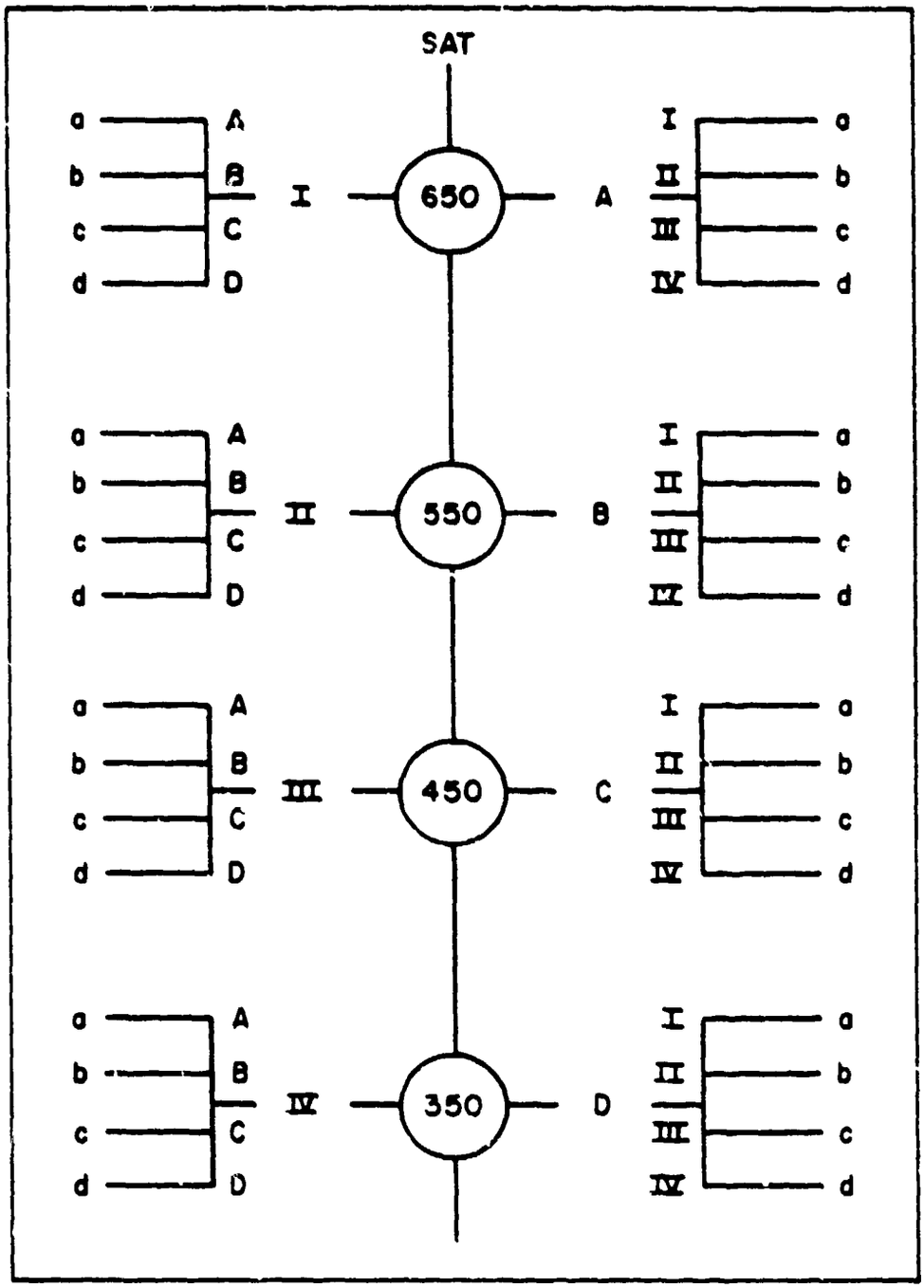

Figure 1. On the left, where SAT scores (350, 450, 550, 650) sort students into colleges (I, II, III, IV), college grades (A, B, C, D) correspond with high school grades (a, b, c, d) but not with SAT scores; on the right, where high school grades sort students into colleges, college grades correspond with SAT scores but not with high school grades.

predictors in the selection process. Though unknown, therefore, these correlations are important enough to estimate. Requiring knowledge of the simple correlations with college of the SAT (.62), FCA ($.62r_{SF}$), and the HSR (.20), developed in the Appendix from published data on the assumption that the left side of Figure 1 rather accurately describes current selection practice, this estimation uses formulas relating the corresponding part and simple correlations to yield $r_{SF} = .62$ and $r_{HF} = .55$.[1] The predictive validities thus increase from .52 to .55 for the HSR and from .41 to .62 for the SAT when college-to-college variation in academic standards is taken into account.

Incremental Validity of the SAT for the Prediction of a Common College Grade-point Criterion

Perhaps the most telling argument against the SAT is that its .06 incremental validity over the HSR, reported by both Slack and Porter and ETS, amounts to only a $100(.58^2 - .52^2)$, or 7, percent increase in the predictable variance of college grades. The

[1] The two part-correlation formulas are

$$.41 = \frac{r_{SF} - .62(.62r_{SF})}{\sqrt{1 - (.62r_{SF})^2}}$$

and

$$.52 = \frac{r_{HF} - .20(.62r_{SF})}{\sqrt{1 - (.62r_{SF})^2}}$$

use of .62 and .55, instead of .41 and .52, to determine the predictive validity of the HSR-SAT combination changes this picture radically, however. According to the formula yielding this validity (.58) from .41 and .52, the correlation between the HSR and the SAT (r_{HS}) is equal to .32 (the same value used in the Appendix to estimate the correlation with college of the HSR).[2] Substitution of .32 for r_{HS} together with .62 for .41 and .55 for .52 in this formula yields a correlation of .72 for the HSR-SAT prediction of first-year college achievement measured on a common grade-point scale. The incremental validity of the SAT for this criterion is thus .72 - .55, or .17, which corresponds to an increase of $100(.72^2 - .55^2)$, or 22, percent in the predictable variance of first-year academic achievement in college. This increase (22 percent) means that the SAT may have over three times the incremental predictive value (7 percent) previously believed.

The numbers used in the foregoing calculations are only estimates of imperfect accuracy. The results certainly reverse the choice, however, between the SAT and the HSR based on their relative predictive validities (.62 for the SAT and .55 for the HSR) or between the HSR-SAT combination and the HSR alone based on the incremental validity of the SAT over the HSR (.17). On these bases, the choice now is clearly the SAT or the HSR-SAT combination, not the HSR alone.

---

[2] This is the multiple-correlation formula

$$.58 = \sqrt{\frac{.41^2 + .52^2 - 2(.41)(.52)r_{HS}}{1 - r_{HS}^2}}$$

Conclusion

Predictive validities of .62 for the SAT and .55 for the HSR make more sense than the corresponding values of .41 and .52. High school teachers do not typically assign grades to predict first-year college performance; however, Educational Testing Service certainly does work to maximize the predictive validity of the SAT. Prediction is the purpose of the SAT, and the predictive validity of .62 indicates that the developers of the SAT have achieved this purpose. The predictive validity of .72 likewise indicates that the HSR-SAT combination is a powerful predictor of first-year academic achievement in college. If college-admissions officers are to abandon the use of the SAT in the student-selection process, therefore, their reason must be something other than that the SAT might have a lower predictive validity than the HSR or that the incremental validity of the SAT might be insufficient to justify its use. In reaching this conclusion, however, this report may only have confirmed what college admissions officers have known from experience for a long time: The SAT is nothing if not valid, and valid it is.

If everyone believed that predictive validity is as important for selection as it is for the SAT, nothing further need be said. Predictive validity is not the only dimension of concern in selection, however. Amid strong cross-currents of values, college-admissions officers must attempt to avoid not only bias but also the appearance of bias in the selection process. Although bias is beyond the scope of this report, values are inescapable. Attendance at one college must have a greater value than attendance at

another college, or the controversy over the SAT would have no meaning. Progress, whether of society or in society, is not the truly relevant value of selective college attendance, however. Prestige notwithstanding, no one college is best for all, and not every college is only for the best. This report has thus come full circle. To say that the SAT or the HSR-SAT combination is valid is to say that it can match students to colleges well. The value of selective attendance at a college--the truly relevant value--depends on this student-college match.

References

Educational Testing Service. Test use and validity. Princeton: Author, 1980.

Ford, S. F., & Campos, S. Summary of validity data from the Admissions Testing Program Validity Study Service. Princeton: College Entrance Examination Board, 1977. (a)

Ford, S. F., & Campos, S. Summary of validity data from the Admissions Testing Program Validity Study Service. Appendix, On further examination: Report of the Advisory Panel on the Scholastic Aptitude Test Score Decline. New York: College Entrance Examination Board, 1977. (b)

Goldman, R. D., & Widawski, M. H. An analysis of types of errors in the selection of minority college students. Journal of Educational Measurement, 1976, 13, 185-200.

Jackson, R. A summary of SAT score statistics for College Board candidates. Appendix, On further examination: Report of the Advisory Panel on the Scholastic Aptitude Test Score Decline. New York: College Entrance Examination Board, 1977. (a)

Jackson, R. Correlations of SAT Scores with high school record. Appendix, On further examination: Report of the Advisory Panel on the Scholastic Aptitude Test Score Decline. New York: College Entrance Examination Board, 1977. (b)

Jackson, R. The Scholastic Aptitude Test: A response to Slack and Porter's "Critical Appraisal." Harvard Educational Review, 1980, 50(3), 382-391.

Lord, F. M., & Novick, M. R. Statistical theories of mental test scores. Reading, Massachusetts: Addison-Wesley, 1968.

Nairn, A., & Associates. The reign of ETS: The corporation that makes up minds. Washington, D.C.: Ralph Nader, 1980.

Slack, W. V., & Porter, D. The Scholastic Aptitude Test: A critical appraisal. Harvard Educational Review, 1980, 50(2), 154-175.

Appendix

On the assumption that the left side of Figure 1 is a rather accurate portrayal of current practice in selection, this Appendix uses published data to estimate the correlations with college of the SAT (r_{SC}), FCA (r_{FC}), and the HSR (r_{HC}) required to determine the predictive validities r_{SF} and r_{HF} from their observed part-correlation counterparts (.41 and .52).

Each correlation with college is an intra-class correlation, a ratio to the population standard deviation of the standard deviation of college means for the variable involved (SAT, HSR, or FCA). These standard deviations are known or directly estimable for the SAT. The population standard deviation for the SAT has remained stable over many years at around 108 (Jackson, 1977a, Table 1); the corresponding median within-college standard deviation, determined to approximate 85 from data published by Ford and Campos (1977b, Table 4), yields as an estimate of the standard deviation of SAT college means $\sqrt{108^2 - 85^2}$, or 67. The value of r_{SC} is thus approximately 67/108, or .62.

The left side of Figure 1 shows that <u>at each SAT level</u> no correlation exists between college and high school or college grades. Insofar as the SAT predominates in the selection process, therefore, within subpopulations of students having equal SAT scores the part-correlation counterparts of the remaining two correlations with college will tend to be equal to zero. The numerators in the formulas for these part correlations should tend,

correspondingly, to be equal to zero: $r_{HC} - r_{HS}r_{SC} = 0$ for the HSR and $r_{FC} - r_{SF}r_{SC} = 0$ for FCA. Substitution of the .62 just estimated for r_{SC} thus yields $r_{HC} = .62r_{HS}$ and $r_{FC} = .62r_{SF}$. According to Jackson (1977b, Table 2), the value of r_{HS} over the years has averaged around .32, and so $r_{HC} = .62(.32)$, or .20, approximately. No further estimation is necessary. Return now to the part-correlation formulas for .41 and .52 (Footnote 1) completes the estimation process.

DISTRIBUTION LIST

	Number of Copies
Defense Documentation Center Cameron Station, Building 5 5010 Duke Street Alexandria, VA 22314	2
Dean of Research Code 012 Naval Postgraduate School Monterey, CA 93940	1
Library (Code 0142) Naval Postgraduate School Monterey, CA 93940	2
Library (Code 54) Naval Postgraduate School Monterey, CA 93940	2
Professor J. K. Arima	1
Professor R. S. Elster	1
Professor J. M. Fremgen	1
Professor C. P. Jones	1
Professor J. D. Senger	1
Professor R. A. Weitzman Code 54 Naval Postgraduate School Monterey, CA 93940	40
Norman Cliff Department of Psychology University of Southern California Los Angeles, CA 90007	1
Bert F. Green, Jr. Department of Psychology The Johns Hopkins University Baltimore, MD 21218	1
Harold Gulliksen Department of Psychology Princeton University Princeton, NJ 08540	1
Alexandra K. Wigdor Associate Study Director Committee on Ability Testing National Research Council 2101 Constitution Avenue, N.W. Washington, DC 20418	1

	Number of Copies
H. Wallace Sinaiko Smithsonian Institution 801 N. Pitt Street Alexandria, VA 22314	1
Glenn L. Bryan (Code 450) Marshall J. Farr (Code 458) Office of Naval Research Arlington, VA 22217	1 1
Eugene Gloye ONR Branch Office 1030 Green Street Pasadena, CA 91101	1
Charles F. Davis ONR Branch Office 536 S. Clark Street Chicago, IL 60605	1
Martin Wiskoff (Code 310) Navy Personnel Research and Development Center San Diego, CA 92152	1
William W. Turnbull President Educational Testing Service Princeton, NJ 08541	10
Fred A. Hargadon Dean of Admissions Stanford University Stanford, CA 94305	1
Richard W. Moll Dean of Admissions University of California Santa Cruz, CA 95064	1
Henry M. Littlefield Headmaster The York School Monterey, CA 93940	1