

Muelder (1966) agrees with this position, and cites, for example, mathematics as requiring extensive and intensive reading, since data are cumulative with knowledge building on other knowledge, and the student must get complete understanding as he progresses in the subject matter. "Because a single word may change the whole meaning in mathematics," says Muelder, "every single word is crucial." This may not be true in social science.

In the field of science, the student is lost if he does not see the organization of information, classification, sequence, explanation of a process, detailed statements of fact, the main idea, details as they relate to the main idea, or is unable to make inferences, and accurate generalizations. In fact, each discipline contains a technical vocabulary which the student must master, if he is to become competent in it.

This review addresses the question: What do we know about the relationship of reading comprehension as measured by standardized test scores to grade point averages of students in engineering, business, mathematics/science, in sociology, and history? Efforts to uncover systematic replications of studies to support theoretical constructs on the predictive validity of standardized reading tests to grade point averages were practically futile. However, the literature contains numerous studies in a combination of disciplines which may or may not duplicate the exact subjects under consideration here.

Grade point average in this review is used as the criterion variable. Both Anastasia (1968) and Lanvin (1966) suggest that academic performance, as indexed by grade point average (GPA), is subject to analagous difficulties, not only because of the subjective criteria that may enter the grading process by different teachers, but because of the uncontrolled sources of variation in or across disciplines. Lanvin (1966) goes a step further to question whether it is meaningful to compare averages of students across disciplines, suggesting that it might be more appropriate to study the determinants of performance within subject matter disciplines. Nevertheless, researchers continue to utilize linear correlation and multiple-regression procedures in their studies.

Citing both the College Ability Test (CAT) and the Scholastic Aptitude Test (SAT) as examples of predictive tests now being subjected to widespread debate and criticism, Bracey (1980), Wickenden (1980), and Stanley (1981), while acknowledging that high school grades have always been a better predictor of academic success in college, suggest that the SAT does in fact predict first year GPA's with the average correlation usually given as .41.

Stanley (1981) of Johns Hopkins bases his faith in the SAT on scores of thousands of precocious youngsters who entered his mathematics programs from junior high school, and who are now—at age 17-18—undertaking Ph.D. degrees at Rutgers, Stanford, Columbia, etc. "The SAT predicts rather well," said Stanley, "which students can move ahead successfully at a fast pace in subject matter too difficult for the typical youth their age." (p. 10)

A report by Slack and Porter (1980) of Harvard casts serious doubt on the foregoing assertions. Their data in eleven studies indicate that the predictive contribution made by the SAT is indeed small, and that superior predictions can be made from a combination of grades and achievement tests. In commenting on the use of tests as predictors of GPA's, Anastasia (1968) contends that no distinction between aptitude and achievement tests can be rigidly applied, since both types are highly verbal in nature. She further states that the STEP, ACT, OSUPT, MAT, and the ITED all sample a combination of general aptitudes and knowledge about subject matter in the major disciplines. Within this general framework, then, the studies reviewed below attempt to answer the question: Do standardized test scores predict GPA's of college students in various disciplines?

WHAT RESEARCH SAYS TO THE DISCIPLINARIAN

EFFIE KAYE ADAMS

East Texan Educators Research Council, Inc.

The noted social scientist, Albert H. Burrows (1961), briefly defines a scholarly discipline as any organized body of knowledge in a given field, and categorizes the scholarly disciplines under two broad headings: Sciences and Non-Sciences. Disciplines under discussion fall into both categories. The interrelatedness of these fields is well known. For example, the sciences of physics and mathematics are drawn upon for the technological discipline of a school of engineering or school of industrial arts; and the sciences of biology and chemistry are drawn upon for materials from which to organize the technology of a school of medicine, nursing, and other healing arts. Similarly, sociology, as a social science, is also a technology as it involves the principles governing group interactions and the effect on the behavior of the individuals and groups participating in them; thus, social service is the applied phase of sociology.

Scholars in the sciences and non-sciences express the belief that there is no substitute for the highly trained disciplinarian, and that in order to become a scholar in any field, one must possess a high level of reading comprehension. Almost four decades ago, Hall and Robinson (1945) reported one of the few analytic studies of the relationship between tests of general reading ability and reading in such disciplines as history, geology, art, and fiction. They identified at the college level the following reading skills: attitude of comprehension accuracy, word meaning, rate of inductive reading, rate for unrelated factors, and chart-reading. Both concluded that reading skills for prose and non-prose were different.

More recently, however, Shepherd (1963) pointed out that reading per se has no content of its own, and must draw from the subject matter disciplines. Therefore, in order to comprehend the materials, one must have a set of reading skills common to all disciplines and also a set of skills peculiar to each discipline.

STUDIES INVOLVING TESTS AS PREDICTORS OF GPA'S

The specific objective of a study undertaken by D'Amico, Bryant, and Pahl (1959) was to determine how well the Multiple Aptitude Test (MAT) correlated with grades in junior college subjects. Test results of 450 students at Flint Junior College and their subsequent grades provided the data for the study. Means, standard deviations, Product-Moment Correlations, and Multiple Correlations were computed. The results showed a range of single validity coefficients varying from $-.01$ to $.73$. Additionally, the Verbal Comprehension Factor contributed significantly to prediction of academic success in Biology, General Business, Psychology, English Composition, Foreign Language, and Drawing; while Test 6, Arithmetic Computation, contributed significantly to the prediction of success in Mathematics, Accounting, Social Sciences, and Business Mathematics. GPA's in each of the subject areas approximated the letter grade of "C," with the standard deviation varying from $.97$ in Drawing to 1.7 in Psychology. Thus, D'Amico and associates (1959) concluded that MAT scores have from moderate to high validity in relation to achievement in junior college subjects.

During the same year, the Iowa Silent Reading Tests, Advanced Form, were among the seven predictors of GPA's of freshmen nursing students at the Trumbull Memorial Hospital in Warren, Ohio in a study by Garrett (1960). Other predictors in the study were: Ohio State University Psychological Tests, Form 21; George Washington University Test Series; Arithmetic Test for Prospective Nurses, Form I; Minnesota Vocational Test for Clerical Workers; MacQuarrie Test for Mechanical Ability; Shipley-Hartford Retreat Scale; and High School GPA's. Multiple Correlations resulted in an 0.641 coefficient for High School Grades, the Iowa Silent Reading Test, the Hunt Arithmetic Test, and the Minnesota Clerical Test as being the most practical predictors in combination. Excluding high school grades, however, the most practical predictors of first year grades in this study include the Ohio State Psychological Test, the Hunt Arithmetic Test, the Iowa Silent Reading Test, and the Shipley-Retreat Scale, which together yielded a multiple R of 0.435 .

Reid, Johnson, Entwisle, and Angers (1962) reported a study in which 410 freshmen at Newark College of Engineering obtained a verbal mean score of 493.3 on the SAT; verbal score of 39.3 on the CAT, Form I-A; and a Composite mean reading score of 28.1 on the Cooperative English Test, C-2. Those entrants obtained a higher mean math score on both the SAT (576.7) and the CAT (42.5) than on the verbal sections. The study failed to show a positive relationship between reading or math scores and GPA's, since 124 or 30 percent of the students were dismissed for unsatisfactory performance (GPA's of -1.25) prior to graduation. In fact, drop-outs at the end of the first semester fell significantly below the entering class mean in mathematics on the SAT, and CT as well as on the entering class mean on the verbal sections of both the CAT and the Cooperative English Test.

A study by Greenwood (1962) utilized the Buffalo Reading Test (BR); the verbal and quantitative sections of the ACE; the SAT; and the Engineering and Physical Science Tests as predictors of GPA's of 444 students who entered the chemical electrical, and mechanical curricula of the Broome Technical Community College, Erie County Technical Institute, and the New York City Community College of Applied Arts and Sciences. The highest correlations for the BR and GPA's was $.60$ and $.70$ for the Broome Technical Electrical Class and the Broome Technical Chemical Class respectively. No other correlations including the reading test were shown. The author recommended that each technical college develop its own predictors, separately.

Lepley (1965) used GPA's of entering freshmen to establish the predictive validity of the Synonym Vocabulary Test, Form I and Form II. The students were divided into four groups:

liberal arts, engineering, science, and non-science. The results indicated that the range of inter-reliabilities between the two forms was from $.82$ to $.90$ with the four groups. In correlating the tests with GPA's, Lepley concluded that the usefulness of the tests was better for liberal arts students (Form I, $.39$; Form II, $.38$) and for science students (Form I, $.46$; Form II, $.51$) than for non-science and engineering students.

A total of 520 students comprised the sample from which Borup (1971) drew data concerning their ACT scores and first semester college GPA's among other variables. Analysis of Variance was computed to determine the amount of relationship between the variables. Comparisons of mean-quarter rankings of male and female students revealed that females exceeded the males in their high school achievement (2.11 vs 1.76), although males scored significantly higher than females on the ACT cumulative and the various sub-scores with the exception of the English portion. Males obtained a lower (1.93) first semester GPA than females (2.31), the F value being 25.69 with significance greater than $.05$. Other variables were studied in this report such as sex and race. However, they are not germane to the present study.

The predictive validity of three new tests with GPA's in selected courses was investigated by ABU-SAYF and George Za'Rour (1975) of the American University of Beirut. Tests were: The English Proficiency (ER), the Quantitative Aptitude Test (AQ), and the Science Proficiency (AS). The participants were 271 freshmen and sophomores enrolled at the University. Results showed that the EP Test predicted grades in English 201 at a higher level of accuracy than other English courses; AQ was found to be a valid predictor of grades in mathematics courses; and Form 886 of the AS Test proved to be a more valid predictor of GPA than was Form 885. Critical ratios between means suggested that science students performed significantly ($p < .01$) better than did the art students. In general, a low to moderate predictive validity for each of the tests was obtained.

Sherrell (1981) investigated the relationship of the N-D Reading Test total mean score and GPA's of 267 students enrolled in three Vocational/Technical Schools in Missouri, and found that as a group, they performed at a total grade equivalent of 10.9 . Students in individual courses earned the following TGE's: Agriculture— 9.5 ; Auto Body— 9.1 ; Auto Mechanics— 10.4 ; Electrical Engineering— 12.3 ; Heating and Air-conditioning— 10.0 ; and Licensed Practical Nursing (LPN)— 13.5 . Of the total sample, 48 or 18% earned A's; 109 or 41% earned B's; 86 or 32% earned C's; 16 or 6% earned D's; and 8 or 3% were dropped (perhaps earning F's). Sherrell complained that all textbooks used by these students had readability levels above the students' reading levels as measured by the N-D, and that the high scores made by students were due to the individualized attention, the use of supplementary reading materials, and teacher demonstrations.

Two final studies were reported by Geer (1981) and Jones (1981) in which different statistical analyses were made to determine the predictive validity of the N-D Tests for GPA's of their majors. In the Geer study (1981) a simple linear regression analysis was performed between Nelson-Denny total mean reading scores and GPA's for 141 LeTourneau College students in Engineering (65), Business (45), and Math/Science (31). Statistical significant correlations were found between GPA's and test scores ($.37$, $.60$, $.66$ respectively) in each discipline at the $<.01$ level, indicating low to moderate predictive validity of the Nelson-Denny Reading Test for GPA.

SUMMARY

This review has been concerned with a small sample of studies undertaken in four decades—from 1959 to 1981,—to assess the relationship of reading comprehension test scores to GPA's of students majoring in Engineering, Business, History,

Sociology, Math/Science. Studies briefly reviewed include both single and multiple batteries of standardized tests as predictors of academic success as determined by grade point averages.

Of the thirteen studies mentioned only six showed moderate to high predictive validity of the tests involved; while seven showed little or no predictive strength for the tests.

Both Alexander (1976) and Walker (1981) concur with Lanvin (1966), Anastasi (1968), Astin (1971), Chansky (1973), and Baird (1979) in the belief that factors other than reading test scores have significant impact on predicting academic success of college students, and that test scores play only a supportive role in the matter of prediction. In fact, Lanvin (1966) and Glaser (1979)—more than a decade—later both expressed the view that efforts to improve predictive and diagnostic validity have run into diminishing returns, and hint that further progress may come only through systemic theorizing about the determinants of academic performance. One implication, as seen by this author, is for replication by reading counselors using identical tests, methodologies, majors, and (as far as possible), similar populations in similar colleges. Until this is done consistently, there is likely to be continued speculations and skepticism in the use of test scores as valid (sole) predictors of academic success at any level.

REFERENCES

- ABU-SAYF and Georga Za'Rour. Predictive Validity of the American University of Beirut Trial Aptitude Battery. *Educational and Psychological Measurements*, 1975, 35, 451-454.
- Adams, Effie Kaye. The Reading Performance of Student Teachers in A Developing Institution, in Schick, George B. and May, Merrill (Eds.) *Junior College and Adult Reading Programs [Expanding Fields]*. Seventeenth Yearbook of the National Reading Conference, 1967, 47-57.
- Alexander, H. T. The Relationship of Reading and Other Factors to Academic Achievement, *Reading World*, 1976, 15, 304-309.
- Anastasi, Anne. *Psychological testing*. (New York: Macmillan Publishers, 1968).
- Astin, Alexander W. *Predicting academic performance in college*. (New York: Free Press, 1971).
- Baird, Leonard. Development of an inventory of documented accomplishments for graduate admissions. (Princeton, New Jersey: Educational Testing Service, 1979).
- Borup, Jerry. The Validity of American College Test for Discerning Potential Academic Achievement Levels—Ethnic and Sex Groups. *Journal of Educational Research*, 1971, 65 (1), 3-5.
- Bracy, Gerald W. The SAT, College Admissions, and the Concept of Talent: Unexamined Myths, Unexplained Perceptions, Needed Explorations. *Phi Delta Kappan*, 1980, 62, (3), 197-198.
- Burrows, Albert H. Science and Scholarship: Interdisciplinary Relationships. *Journal of Educational Sociology*, 1961, 35, 181-188.
- Chansky, Norman. Resolving the Grading Problem. *Educational Forum*, 1973, 37 (2), 189-194.
- D'Amico, Louis A., Bryant, J. Howard, And Prahl, Marie. The Relationship Between MAT Scores and Achievement in Junior College Subjects. *Educational and Psychological Measurements*, 1959, 19 (4), 611-614.
- Garbarino, James. The Meaning and Implications of School Success. *Educational Forum*, 1976, 40, 157-167.
- Garrett, Wiley S. Prediction of Academic Success in a School of Nursing. *Personnel and Guidance Journal*, 1960, 38, 500-503.
- Geer, Carolie. Predicting Grade Point Averages of Majors in Engineering, Business, Science/Mathematics Through the Use of Reading Comprehension Tests. Paper presented at the second annual conference of the American Reading Forum, Sarasota, Florida, December 12, 1981.
- Geis, Lynna and Carney, Myrna L. Reading ability, college grades, and attrition: A two year study. in Michael Kamil and Alden Moe (eds.) 28th Yearbook, National Reading Conference, 1979, 163-167.
- Glazer, Robert. Trends and Research Questions in Psychological Research on Learning and Schooling. *Educational Researcher*, 1979, 8, (10), 6-13.
- Greenwood, Robert L. The prediction of academic success in the Technical curricula of community colleges: An investigation of the prediction of academic success in the chemical, electrical, and mechanical curricula of three community colleges in New York stat. Doctoral Dissertation, New York University, 1962. D.A. 23:898-899.
- Hall, William and Robinson, Francis. An analytical approach to the study of Reading Skills. *Journal of Educational Psychology*, 1945, 36, 429-442.
- Hansmeier, Thomas. The Iowa Tests of Educational Development as predictors of college achievement. *Journal of Educational and Psychological Measurement*, 1960, 20 (4), 843-845.
- Jones, Rube Jo. Comparison of reading comprehension test scores and grade point averages of majors in history and sociology. Paper presented at the second annual conference of the American Reading Forum, Sarasota, Florida, December 12, 1981.
- Lanvin, David E. The prediction of academic performance. (New York: The Russell Sage Foundation, 1965).
- Lepley, William. Delayed validation for the Synonym Vocabulary Test, Form I, Form II. *Journal of Psychology*, 1965, 59, 329-331.
- Muelder, Richard. Helping students read mathematics. In H. Alan Robinson and Sidney Rauch (eds) *Perspectives in Reading in the High School Classroom*, No. 6. (Newark, Delaware: International Reading Association, 1966).
- Reid, John, Johnson, A.P., Entwisle, Frank, and Angers, William P. A Four Year Study of the Characteristics of Engineering Students. *Personnel and Guidance Journal*, 1962, 41, 38-43.
- Shepherd, David. Reading in the content areas: Instruction and application. In Clay Ketcham (ed) *Proceedings of the College Reading Association*, 1968, 9, 43-47.
- Sherrell, Eugene. Long-Neglected factor: Reading ability of vocational students. *Phi Delta Kappan*, 1981, 63 (2), 140-141.
- Slack, Warner and Porter, Douglas. The scholastic aptitude test: A critical appraisal. *Harvard Educational Review*, 1980, May, 154-75.
- Stanley, Julian C. Using SAT to find youths who reason extremely well mathematically: A report on the Johns Hopkins University Mathematics program. *Human Intelligence International Newsletter*, July/October, 1981, p. 10.
- Walker, Mary Margaret. Different Strokes: Raising the level of expectations of college students as a factor in improving reading comprehension scores and grade point averages in various disciplines. Paper presented at the second annual conference of the American Reading Forum, Sarasota, Florida, December 12, 1981.
- Wickenden, James. How the SAT's are used—and sometimes abused. *Chronicle of Higher Education*, September 29, 1980, p. 56.