PREDICTING SCHOLASTIC SUCCESS

JOSEPH P. ROBERTS

George Peabody College for Teachers, Nashville, Tennessee

THE PROBLEM OF PREDICTION

The prediction of human behavior is, at best, hazardous and difficult. Some of the simpler elements of human behavior can be predicted with considerable confidence because we are able to control or measure the major variables affecting such behavior. But as the number and complexity of the variables affecting behavior increase, the confidence with which we can predict behavior decreases sharply. This is due to the difficulty of identifying and defining some of the more complex variables, and to the difficulty of controlling and measuring many of the variables that have been identified and defined. A good example of a form of behavior that is very difficult to predict is, for want of a better term, 'scholastic behavior.' The variables that affect scholastic behavior are fantastically complex, and, with the increasing emphasis on planning educational programs to meet individual needs, we may see more clearly just how complex these variables are.

MEASURES THAT HAVE PREDICTIVE VALUE

In spite of the obvious difficulties, educators and psychologists have developed techniques for predicting, in useful measure, the chances of a given individual's attaining a given measure of scholastic success. This paper is concerned with an appraisal of these techniques with a view toward improving and supplementing them. It is neither practical nor necessary to review all the tests and measures that have been used to predict scholastic success. The traits that are known to have predictive value are relatively few. Intelligence, as measured by certain tests, is, perhaps, the best single predictor. Previous scholastic success is, of course, a good predictor, and is probably relied upon more heavily by college admissions boards than any other measure. There are several limitations to this measure, but a discussion of them is outside the scope of this paper. Achievement tests are also useful as predictors of scholastic success, but they reflect differences in previous experience as well as scholastic aptitude. Good scholastic aptitude or intelligence tests do not, theoretically, give weight to previous specialized training. They should not, therefore, penalize the student with a poor scholastic background or give undue credit to the student with a superior scholastic background. They do have some bias, however, because the best of them are far from perfect. Personality tests are generally inferior to scholastic aptitude tests for prediction. The motivation to succeed scholastically is an important variable in the determination of scholastic success,

but no valid way of measuring such motivation has as yet been discovered.

Although they are not, at present, any better predictors of scholastic success than achievement tests or records of past grades, scholastic aptitude tests seem to be more promising than any other single predictor. Perhaps the term 'aptitude' should be clarified. An aptitude may be defined as potential ability in a specific area of endeavor. General intelligence may then be thought of as a weighted pattern of aptitudes. This is not an adequate definition of general intelligence, but it indicates that aptitudes are specific.

Scholastic aptitude may then be roughly defined as that "part" of general intelligence which is concerned with the functions of the process of formal education. Several tests have been devised to measure this aptitude, and one of the more widely used of these will be discussed in some detail.

TABLE 1. *Coefficients of correlation between grade averages and ACEPE scores at the University of Chattanooga (Mean No. of graded semester hours—96.4) (Robert's, 1951)*

ENTITIES	N	T WITH GRADES	Q WITH GRADES	L WITH GRADES	Q & L WITH GRADES
Total sample......	281	.56	.40	.55	.56
		.04*	.05*	.04*	
Q majors.........	85		.12		
			.11*		
L majors..........	81			.55	
				.08*	
Economics and Commerce majors	86		.31	.45	
			.10*	.09*	

*Standard error of r

Predicted grade $= .145$ (Q score) $\pm .47$ (L score) (Z score form).

Standard error of estimate $= \pm .83$ (Z score form).

THE AMERICAN COUNCIL ON EDUCATION PSYCHOLOGICAL EXAMINATION

The American Council on Education Psychological Examination, which may be abbreviated *ACEPE,* is an outgrowth of L. L. Thurstone's work on the factor analysis of intelligence, in which he attempts to show that intelligence is made up of independent abilities. The *ACEPE* is composed of six subtests which are based, in part, on these so-called primary mental abilities. Three of the subtests are grouped under what is called quantitative or Q ability, and three under what is called linguistic or L ability. The raw scores for the Q and L sections are added together to obtain a total or T score. According to the *ACEPE Manual of Instructions,* Q scores should be better predictors in the more quantitative fields such as mathematics, engineering, physics, chemistry, etc., and L scores should be better for the

more linguistically loaded fields such as the social studies, history, education, etc.

A Review of Some Representative Studies Involving the ACEPE

Studies that have attempted to measure the differential predictive value of the *ACEPE* differ markedly in their results. Table 1 shows the results of such a study that was carried out by the writer at the University of Chattanooga. As will be shown, the correlations found in this study agree rather closely with those found in other similar studies. In most studies of this nature, L scores correlate higher with overall grade averages than do Q scores, and, in many studies, L scores correlate virtually as high with overall grade averages as do T scores. In the study cited, the correlation between L scores and overall grade avarages was not significantly lower than the multiple correlation of Q and L with grades.

Table 1 gives correlations computed from Q scores and grades of "Q majors," and from L scores and grades of "L majors." These represent divisions of the total sample according to whether the student's major field was predominantly quantitative or linguistic as outlined by the *ACEPE Manual of Instructions*. In dividing the sample into so-called Q and L major fields of study, the choices were necessarily arbitrary and open to criticism. The main difficulty lies in the fact that all fields of study include both Q and L subject matter in varying and essentially indeterminate proportions. Obviously, such arbitrary selections tend to invalidate the results, and further analysis of the pattern of abilities required for each field of study is indicated.

In another study, in California, Andrew MacPhail (1942) concludes that there is no significant differential predictive value of Q and L scores. That is, in his study, Q scores predicted as well as L scores, but no better. In a survey of findings during the years immediately preceding 1943, Glenn Durflinger (1943) found that correlations between the *ACEPE* and grades ranged from .17 to .81, with the median between .40 and .50. Generally, Q scores seem to be more erratic than L scores as predictors, which is probably due in part to the presence of an L factor in virtually all subject matter, while a Q factor is present to a measurable extent only in certain subject matter. Robert S. Morrow (1941) found correlations of .57 between Q and L, .82 between Q and T, and .94 between L and T. In the study at Chattanooga cited above, the correlation between Q and L scores was .54.

In another study in California, Hugh S. Brown (1950) found that Q scores are no better than L scores as predictors of success in quantitative subjects, and that neither was very good (.33 and .30). However, in linguistic subjects, L scores were as effective as predictors as any other test that had been tried (.54).

Essentially the same trends were found by Berdie, Dressel, and Kelso (1951) in a co-operative study involving four state universities, four private universities, two state colleges, two state teachers colleges, and one technical institute. They also found, as did Durflinger,

that predictions for women seem to be somewhat better than for men. Durflinger found, too, that regression equations work better when set up by sex. Osborne, Sanders, and Greene (1950) also found that predictions for women were more accurate, and that first quarter grades were better predictors than *ACEPE* scores. As in the study at Chattanooga, W. L. Wallace (1949) found that L scores were better predictors of overall grade averages than were Q scores.

Any review of previous studies will show that the *ACEPE* has predictive value, and that Q, L, and T scores have different values for different schools. The *ACEPE* can be a very useful predictor if it is properly validated on the local population with which it is to be used. Some schools have found decidedly significant differential predictive values for Q and L scores, and these values are, to some extent, functions of the nature of the populations, teaching and grading methods, etc.

DESIGNING THE TESTING PROGRAM

Some of the factors to be considered in setting up a testing program for predicting scholastic success are: (1) The validity of the independent variable; does the test measure any important aptitude that is required to master academic subjects? (2) The validity of the criterion; does the criterion, usually grade average, accurately reflect scholastic success? That is, does grade average reflect the proper proportion of the independent variable (the predictor test)? If scholastic success is defined in terms of grade average without regard for the factors that contribute to grade average, as is often the case, the predictive efficiency of scholastic aptitude tests will be interpreted as being lower than is actually the case. This is true because the errors inherent in grading will not be taken into account. Unfortunately, accurate criteria are very hard to come by, and the best that can be done with grade averages is to (3) consider their validity as a function of the number of courses and the number of teachers that contribute to them. Teacher bias, temporary emotional upsets, illness, etc., will tend to have less effect on grade averages as the number of courses and the number of teachers increase. However, this is not an unmixed blessing. Some studies show that *ACEPE* loses predictive efficiency as the time from testing increases. The fourth factor to be considered in designing a testing program for predicting scholastic success is the grouping of independent variables. A properly selected battery of tests is almost always a better predictor than any single test. Durflinger found the median multiple correlation to be between .60 and .70, and rarely higher than .80 regardless of the independent variables used. He also found that an intelligence test, a good achievement test, and high school grades usually yield the highest multiple correlations. Rank in high school class is usually a better predictor than high school grade average.

Most scholastic aptitude tests are concerned with predicting success in higher education. The compulsory nature of education through the high school level, plus the necessity for colleges and universities to restrict enrollment to the more able students, accounts for this

emphasis. Thus, all of the studies cited here have been concerned with predicting scholastic success on the college level.

SUMMARY

1. Scholastic aptitude tests have some advantages over other measures as predictors of scholastic success, and are more promising for future improvements in such prediction.

2. The *ACEPE* is very useful as a predictor of scholastic success, particularly when it is standardized on the local population with which it is to be used.

3. Ideally, a testing program set up to predict scholastic success should include a careful analysis of the independent and dependent variables, and some assurance of their stability and adequacy. These latter considerations suggest a longitudinal measure with respect to the dependent variable, and several measures for the independent variable.

LITERATURE CITED

American Council on Education. 1949. *Psychological Examination, Manual of Instructions.* Cooperative Test Division, Educational Testing Service, Princeton, New Jersey.

Berdie, Ralph, Paul Dressel, and Paul Kelso. 1951. Relative validity of the Q and L scores of the ACE Psychological Examination. *Educ. and Psychol. Measurement,* 11 (4):803-812.

Brown, Hugh S. 1950. Differential prediction by the A. C. E. *Jour. Educ. Research,* 116-121, Oct.

Durflinger, Glenn W. 1943. The prediction of college success—a summary of recent findings. *Amer. Assoc. Collegiate Registrars,* 19:68-78.

Morrow, Robert S. 1941. An experimental analysis of the theory of independent abilities. *Jour. Educ. Psychol.* 32:495-511.

MacPhail, Andrew H. 1942. Q and L scores on the A. C. E. Psychological Examination. *School and Society,* 56:248-251.

Osborne, R. Travis, Wilma B. Sanders, and James E. Greene. 1950. The differential prediction of college marks by A. C. E. scores. *Jour. Educ. Research,* 107-115, Oct.

Roberts, Joseph P. 1951. An investigation of the differential predictive value of the American Council on Education Psychological Examination. Unpublished honors thesis, University of Chattanooga.

Wallace, W. L. 1949. Differential predictive value of the American Council on Education Psychological Examination. *School and Society,* 70:23-25.