Accuracy of Clinical Judgments of Attrition and Survival of Students in Engineering Training

Albert B. Truesdell
Iowa State College
Accuracy of Clinical Judgments of Attrition and Survival of Students in Engineering Training

By Albert B. Truesdell

Clinical tests have received widespread acceptance as tools for predicting and assessing human potentialities. A casual review of the testing area reveals the existence of a large number and variety of such instruments. In the use of tests, the counselor is the common denominator—he is the one who clarifies, interprets and makes recommendations to the counselee.

Sherman (1) found that experienced clinicians, using Kuder profiles, could more accurately predict curricular appropriateness in the physical and biological sciences than in English. Appropriateness is defined as grade point average. If a student gets a higher grade point average in English than chemistry, English is considered more appropriate than chemistry. Berdie's (2) study indicates that there is a high reliability coefficient between the counselor’s original statement of vocational appropriateness and a later prediction of appropriateness for the same counselee based on unidentified case folders. Goche (3) found significant differences between some of the sub-scales of the Kuder Preference Record (form CH), Guilford Martin Temperament Traits (GAMIN), Guilford Temperament Traits (STDCR) and first quarter college grade point average in the statistical prediction of attrition and survival of students in engineering training. A discriminant function was developed to determine the contribution of each of the significant variables to the prediction equation of students classified as survival. This point will be brought out more fully later in this study.

The purpose of this study is to determine the accuracy of subjective judgments in making predictions of success and failure of students in engineering training.

Method

During the fall of 1950, 314 male freshmen were registered in the various engineering curricula at Iowa State College. At that time they were administered the customary freshman battery of tests, plus additional interest and temperament tests. By the summer of 1953, 52 per cent of these students had dropped out of
engineering training. The test data for the Kuder Preference Record (form CH), Guilford Temperament Scale (STDCR), Guilford Martin Temperament Scale (GAMIN) and first quarter college grade point averages for the 314 students were analysed to determine their possible contribution to the discrimination between attrition and survival students. For the purpose of this paper survival is defined as those students who did succeed in engineering training; whereas attrition is defined as those who did not succeed. The discriminate function developed by Goche on this data to determine the contribution of each variable in the prediction equation for survival, reveal six to be significant. They are, in order of predictive value, (1) first quarter college grade point average, (2) Kuder—Computational, (3) Guilford—Ascendence-Submission, (4) Kuder—Scientific, (5) Guilford—Rhythemia and (6) Guilford—Social Introversion-Extroversion.

One hundred random subjects were drawn from the original group of 314 engineering students: 52 from the attrition group and 48 from the survival group. For the present study, a combined profile of the Kuder scores, Guilford Temperament Trait scores and first quarter college grade point average was developed on each subject.

These profiles were presented to each of nine judges independently. Each judge was asked to make a prediction of attrition or survival for each student on the basis of these data.

For the purpose of examining the possible influence of varying amounts of clinical experience upon judgments, each judge was classified into one of three groups according to his experience in dealing with clinical data. Three judges were classified as “A”, three as “B” and three as “C”, experience being the greatest for the “A” group and least for the “C” group. The test scores and grade point average for each student were also judged in the light of a multiple cutting scores developed from the discriminant function. Nine subjective judges and one objective judge were involved in this study. The score for each judge was analysed by means of the phi coefficient of correlation.

Chi square analysis of inter-group differences was rendered invalid because each judge made prediction on the same profiles which produced correlated results. An approximation of the inter-group differences was then needed. According to Cochran (4) if the chi square value for correlated or matched data approaches the significance level of an F value in an analysis of variance, quantitative analysis can be used to approximate qualitative results.
Results

An analysis of variance for the data was performed. Each right prediction was given a value of 1, each wrong prediction a value of 0. This produced an F significant at the 1 per cent level of confidence, indicating that significant differences between judges’ ability to predict attrition and survival do exist.

Cochran’s chi square test was applied to the data yielding a value significant at the 1 per cent level of confidence. It was then assumed that a quantitative estimate of the qualitative results could justifiably be made.

The number of right predictions of nine out of ten judges correlated at the 1 per cent level of significance with the actual number of attrition and survival cases in this study. The predictions of one judge, in group “A”, did not correlate significantly with the objective criterion.

The mean right predictions for each group are: “A”, 68, “B”, 70.66; “C”, 73.33. The discriminant function successfully predicted 75 right out of a possible 100.

Contrasts were made between the groups and the discriminant function to determine which was the most accurate in the prediction of attrition and survival. These intergroup comparisons indicate the existence of a significant difference at the 5 per cent level of confidence between group “A” and group “C”. A difference significant at the 5 per cent level (approaching significance at the 1 per cent level) exists between group “A” and the discriminant function. It is thought that this difference is due to the one judge in group “A” who accurately predicted only 56 out of a possible 100 profiles. If this judge be considered a typical of the rest of the judges and dropped out of the analysis, the mean number of right predictions for group “A” is raised to 74. This would eliminate any significant inter-group differences.

When the groups of judges and the discriminant function were ranked in order of accuracy of predictions, the rank was (1) discriminant function, (2) group “C”, (3) group “B” and (4) group “A”. By eliminating the one judge from group “A”, the ranking was (1) discriminant function, (2) group “A”, (3) group “C” and (4) group “B”.

Discussion

It is felt that inter-group differences would be increased if a larger number of judges were in each group. As it is, with only three judges in each group, the results concerning the inter-group differences must be interpreted with reservations.
A polling of the judges indicated that greater emphasis was placed on grade point average than any other index in their predictions. Inasmuch as the use of grade point average is not restricted to the clinical situation, it is reasonable to assume that the results would not reflect the effect of clinical experience. It can be hypothesized that a similar study using only clinical test data would accentuate the differences between the groups of judges.

The assigning of judges into groups must be interpreted as arbitrary at best. Amount of education and degree of familiarity with clinical data was not held constant in this study. Judges were merely assigned to groups according to the proportion of time they were devoting to clinical work at the time of this study.

This indicates that judges can accurately predict survival and attrition of students in engineering training on the basis of the information presented in this study. It must be noted that significant differences between the judges' ability to make accurate predictions do exist and must be taken into consideration before these results can be generalized.

Bibliography