

A NOTE ON THE PREDICTION CEILING

By BRIAN BOLTON, Ph.D.

In a recent paper Rundquist (1969) pointed out that validity coefficients in industrial and personnel psychology seldom exceed 0.50 and show no time trend toward improvement. He presented a systems analysis strategy for discovering the mediating processes in learning and performing tasks and ultimately raising validity coefficients above this apparent "prediction ceiling."

It is assumed, in theory at least, that the validity of a test procedure is limited only by the unreliability of the test and the criterion. The purpose of this paper is to suggest that factors inherent in the test situation operate to reduce validity coefficients. Stated somewhat differently, true ability (optimal performance under non-test situation conditions) may not be displayed in the testing situation. For example, work samples or job replicas should predict criterion behavior very well; lack of validity may be due to contaminating factors on the job (e.g., motivation, interpersonal relations, etc.), but may also be due to extraneous factors in the predictor situation (e.g., test anxiety, lack of test wiseness, etc.). To speak of factors in the test situation is misleading; the test situation only serves to provoke test anxiety or to penalize persons who are not test wise.

The data reported in this study were collected in conjunction with a coaching program for deaf persons conducted jointly by the Illinois Division of Vocational Rehabilitation (DVR) and the Chicago Jewish Vocational Service (CJVS). Deaf clients were referred by DVR counselors to the coaching program which was organized and administered by the CJVS. The goal of the two week program was preparation for the Clerk Carrier Examination for post office employment (U. S. Civil Service Commission, 1969).

Sample.

The subjects of this study were 49 profoundly deaf persons. They ranged in age from 20 to 55 with an average age of 34, were all in the average intelligence range, and most had some high school education. All subjects were currently employed or had been employed in the recent past; thus their vocational status was "upgrading."

The 49 subjects were about equally divided among six coaching classes. Each class was conducted by an instructor who was fluent in the manual language of deaf communication. The eight or nine sessions which comprised the coaching program were devoted to practice exercises and sample examinations which were preceded by detailed instructions and demonstrations and followed by an analysis of individual errors and problems.

Instruments.

The "predictor" instrument used in this study was the Minnesota Clerical Test (Andrew & Paterson, 1969). The MCT is a test of speed and accuracy in performing tasks related to clerical work. It consists of two parts, number checking and name checking. The examinee's task is to check identical pairs of numbers and names as quickly as possible.

Anastasi (1968, p. 364) states that the reported retest reliabilities which are in the 0.70's and 0.80's for each of the two parts are under-estimates. Of importance to the current study is the evidence that scores on a second administration are appreciably more reliable than scores on the first administration (Andrew & Paterson, 1959). Furthermore, the authors indicate that the two subtests, number checking and name checking, correlate 0.70 on the average (the subtests correlated 0.70 and 0.68 for the two MCT administrations of the current study). Correlations of this magnitude justify summing the subtest scores for a more reliable total score. The reliability of the total scores for the second administration is estimated to be at least 0.90.

The "criterion" measure was the Address Checking (AC) subtest of the Clerk Carrier Examination (U. S. Civil Service Commission, 1969). It consists of 95 pairs of addresses (number and street, or city, or geographical location) which are exactly alike or different in some respect. The examinee's task is, again, to check the similar pairs as quickly as possible. AC is scored the same as the

MCT—rights minus wrongs. The AC subtest is most certainly less reliable than the MCT because it is shorter (six minute time limit as compared to a 15 minute time limit for the MCT total score). The reliability of AC is estimated from the Spearman-Brown formula using the MCT data to be at least 0.74. (It is worth noting as a point of interest that Anastasi (1968, p. 363) attributes clerical matching ability to a perceptual speed factor and Guilford (1967, p. 405) indicates that she is in error in doing so.)

Methodology.

The logic of the design entails the prediction of test performance (the "criterion" measure, AC) from non-test situation performance (the "predictor" measure, MCT). The predictor and criterion measures require the same type of ability; in fact, the tasks are virtually identical. Furthermore, the subjects had practiced on sample AC tests in the coaching sessions. Thus, the "validity" coefficient should approach the geometric mean of the reliabilities of the two measures, which is 0.81. (This is the maximum validity coefficient which could be obtained, given the reliabilities of the measures.) The extent to which this result is not realized will reflect the influence of test situational factors which reduce the validity of test procedures.

The MCT was administered twice during the first week of the coaching program. While data is reported below for both administrations, the scores from the second administration constituted the "predictor" variable. Approximately two weeks later the subjects took the Clerk Carrier Examination at the post office. A deaf proctor was present at the examination. (Scores were kindly supplied by the Civil Service Examination Board in Chicago.)

Results.

The second administration of the MCT recorded a large improvement in performance for this sample of 49 deaf persons. The mean total scores were 188 and 223 on the first and second administrations respectively, an average gain of 35 points. The magnitude of this gain is apparent when compared to the standard deviation of 49 for either administration. When compared to industrial applicant norms (Andrew & Paterson, 1959) this gain represents an improvement from approximately the 25th percentile to the median. The correlation of 0.82 between the first and second administration scores indicates that the size of gain was in-

dependent of initial score; that is, there was a consistent benefit from practice on the first administration.

The correlation between the MCT scores (second administration) and the AC scores for the 49 deaf subjects was 0.58. The score distributions were normal, the scatter plot was rectilinear in form, and there was no restriction of range.

Discussion.

The obtained "validity" coefficient of 0.58 is somewhat lower than that which might be expected based on the statistics presented in previous sections. In fact, this would be regarded as just a good validity coefficient in an industrial setting involving routine assembly work.

The major difference between the prediction and criterion measurements in this study was the test situation atmosphere: the Address Checking test was completed under "real" test conditions. Given the subject's previous practice and familiarity with the test task, it seems reasonable to implicate test anxiety as the major factor contributing to the reduced validity coefficient. This suggestion becomes more plausible when the subject sample of deaf persons is properly considered as a disadvantaged group. In discussing techniques for alleviating test anxiety with disadvantaged groups, Doppelt and Bennett (1967) suggest practice test taking or the liberal granting of retests. The coaching program described above should have helped in this regard, but without a control group, the benefits derived can only be conjectured. It can be tentatively concluded from this study that situational factors associated with test taking contribute to the establishment of the prediction ceiling, at least with disadvantaged groups.

REFERENCES

- Anastasi, A. *Psychological Testing*, (Third Edition) New York: Macmillan, 1968.
- Andrew, D. M., and Paterson, D. G. *Manual for the Minnesota Clerical Test (Revised)* New York: The Psychological Corporation, 1959.
- Doppelt, J. E., and Bennett, G. K. Testing job applicants from disadvantaged groups. (Test Service Bulletin No. 57) New York: The Psychological Corporation, 1967.
- Guilford, J. P. *The Nature of Human Intelligence*. New York: McGraw-Hill, 1967.
- Rundquist, E. A. The prediction ceiling. *Personnel Psychology*, 1969, 22, 109-116.
- United States Civil Service Commission. *Prepare yourself for the Postal Examination*. (AN 2450-R) Washington: U. S. Government Printing Office, 1969.