

STANLEY S. BLANK,
University of British Columbia

AN EXAMINATION OF THE USEFULNESS OF VARIOUS
PSYCHOLOGICAL INSTRUMENTS FOR PREDICTING
DEPARTMENT MANAGERS' RATINGS OF
CLERICAL SALES PERSONNEL

Information concerning the effectiveness of personnel testing programmes for predicting success in department-store selling is sparse at best. While certain psychometric devices have been found to have some little predictive validity for certain kinds of selling functions, the results are far from impressive or reassuring. Most of the reported studies on personnel selection have tended to use supervisory ratings as criterion variables. Indeed, as Guion (1965) points out, of a group of studies done between 1950 and 1955, 81% relied on some kind of rating as the criterion measure. (p. 96)

Since so little attempt has been made to assess the ability of psychological tests to predict to assessment of performance in a department-store setting, one can seriously question the usefulness of these procedures in the selection and placement of department-store personnel. The present study is an attempt to examine the effectiveness of personnel testing as a selection procedure for department-store sales staff.

PROCEDURE

The subjects were randomly selected from sales personnel in four Toronto-area branches of a Canadian department store. Following selection, the participants were given orientation to the testing programme as an experimental procedure designed to validate a new personnel selection procedure. They were then requested to report to the testing area in each store in groups of from 6-10 persons for a one-day testing session.

The tests were administered by personnel clerks specially trained by the investigator, and all followed the same order of test administration, and all used the same standard directions provided in the test manuals. In total, a 6-test battery was administered to 268 subjects.

The following tests comprised the predictor battery and were administered in the order listed: Gordon Personal Profile; Gordon Personal Inventory; Bruce's Sales Comprehension Test; Minnesota Clerical Test; Kuder Preference Inventory—Vocational; and the Wonderlic Personnel Test.

Two rating scales were constructed by the investigator and were used as criterion measures. These were the "Department Relations Scale" and the "Customer Relations Scale." Both scales were completed by the department managers of the personnel concerned after several hours of training seminars conducted by the author.

RESULTS

The statistical technique used in the present study is that of multiple linear regression. (Bottenberg & Ward, 1963) All analysis of variance and

covariance designs are subsumed under the more general multiple linear regression approach. The model assumed is a linear one having the form:

$$\hat{Y}_1 = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n$$

where \hat{Y}_1 = the predicted criterion score

$x_i = 1, n$ — the set of predictor variables

$a_i = 1, n$ — the set of predictor weights so chosen as to minimize prediction errors

a_0 — an additive constant.

The set of predictors may consist of both categorical and continuous predictors.

The basic plan determines weights for the a 's such that the error sum of squares (ESS), defined as $ESS = \sum_1^n (Y - \hat{Y}_1)^2$, is minimized. That is, weights are so chosen that the sum of squared differences between predicted (\hat{Y}_1) and obtained (Y) criterion scores is minimized.

The goodness of fit between \hat{Y}_1 scores and Y scores is described by the product-moment correlation (R) between these scores. The square of this correlation (RSQ) is the proportion of the criterion variance accounted for by the best weighted linear combination of predictor variables.

Hypothesis testing using this approach involves a comparison between the RSQ yielded by the "full" model which contains all the predictor variables (R_1^2) and the "restricted" model (R_2^2). The second model is restricted in the sense that it does not contain the predictor variable whose effect, in the presence of the other predictors one wishes to assess. R_2^2 is less than R_1^2 . The F test used to evaluate the independent contribution of the dropped variable is:

$$F = \frac{(R_1^2 - R_2^2) / df_1}{(1 - R_1^2) / df_2}$$

where df_1 = the difference between the number of independent weights in the full and restricted model and

df_2 = the difference between sample size and the number of independent weights in the full model.

Analysis indicated that only one scale of the Kuder Preference Inventory—Vocational (K9) reached the required level of significance ($p = .012$) (See Table 1). However, two other predictor variables (Gordon Personal Inventory Scale C ($p = .072$) and the Wonderlic Personnel Test ($p = .098$)) reached a high enough level of significance to indicate the usefulness of further study of these two variables and their possible effectiveness as predictors of sales performance.

While none of the remaining predictors accounted for any significant degree of criterion variance, examination of the correlation matrix (Figure 1) reveals a number of interesting significant correlations.

<i>Variable</i>	<i>df</i>	<i>F</i>	<i>Probability of F</i>
Years in Co.	1/238	0.69	0.41
Gordon PI (C)	1/238	3.28	0.07
Gordon PI (O)	1/238	0.00	1.00
Gordon PI (P)	1/238	0.03	0.86
Gordon PI (V)	1/238	0.99	0.32
Gordon PP (A)	1/238	0.99	0.32
Gordon PP (R)	1/238	0.70	0.40
Gordon PP (E)	1/238	2.23	0.14
Gordon PP (S)	1/238	0.89	0.35
Sales Comp.	1/238	0.07	0.79
Wonderlic P.	1/238	2.77	0.09
Minn Clerical 1.	1/238	1.69	0.19
Minn Clerical 2.	1/238	1.19	0.28
Kuder 0	1/238	0.02	0.89
Kuder 1	1/238	0.00	1.00
Kuder 2	1/238	1.99	0.16
Kuder 3	1/238	0.00	1.00
Kuder 4	1/238	1.65	0.20
Kuder 5	1/238	1.65	0.20
Kuder 6	1/238	0.01	0.95
Kuder 7	1/238	0.12	0.73
Kuder 8	1/238	0.00	1.00
Kuder 9	1/238	6.32	0.01

Table 1. F ratios, degrees of freedom, and probabilities for selected predictor variables.

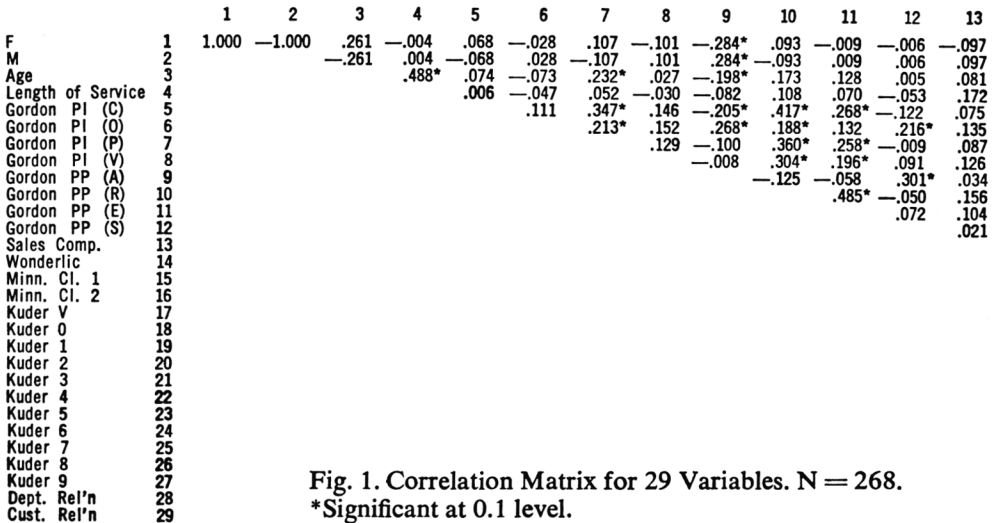


Fig. 1. Correlation Matrix for 29 Variables. N = 268. *Significant at 0.1 level.

agers. Even fairly extensive training periods fail to overcome such differences and point once again to the complete inappropriateness of ratings of employee performance as criteria for establishing the predictive validity of personnel testing programmes.

The present study strongly suggests that as long as department managers' evaluation of employee performance is the essential criterion, the use of group tests of personality and aptitude for the selection of department-store personnel is at best of questionable value. Further, the need for additional and more objective criteria is obvious. This in turn implies the need for developing multiple criteria rather than attempting to use a single overall criterion of sales performance.

REFERENCES

Bottenberg, R. A., & Ward, J. H., Jr. *Applied multiple linear regression*. Washington, D.C.: U.S. Department of Commerce, Clearinghouse for Federal, Scientific and Technical Information. Technical Documentary Report PRL-TDR-63-6, 1963.

Guion, R. M. *Personnel Testing*. New York: McGraw-Hill, 1965.

UN EXAMEN DE L'UTILITE DE DIVERS INSTRUMENTS PSYCHOLOGIQUES POUR PREDIRE LES CLASSEMENTS DES EMPLOYES A LA VENTE PAR LEURS CHEFS DE SERVICE

Cette étude a été préparée de telle manière que l'on puisse s'en servir pour évaluer la mesure dans laquelle certaines épreuves psychologiques permettraient de pronostiquer les résultats de l'évaluation de l'efficacité de leur personnel à la vente par les chefs de départements dans les grands magasins.

On a soumis deux cent soixante huit sujets, vendeurs dans quatre succursales de grands magasins dans la région de Toronto, à un test comprenant six épreuves psychologiques au cours d'une session diurne. Puis, ayant obtenu de leurs chefs de départements les classements d'efficacité de ces sujets, on a étudié les épreuves en tant que pronostics de ces classements. D'après les résultats de cette étude, il est peu probable que l'on puisse, tant que l'on se sert d'un système de classement comme critère d'efficacité, pronostiquer, par moyens d'épreuves d'habileté et de personnalité subies en groupe, les évaluations de l'efficacité des vendeurs par leurs chefs de départements.

Il est donc suggéré que l'on essaie de trouver des critères plus objectifs, et que des critères multiples plutôt que singuliers soient requis dans tout programme effectif d'évaluation de personnel.