The Validation of a Test Battery and Biographical Data Used to Select Key-Punch Operators

Michael F. O’Toole

Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/masters_theses

Part of the Industrial and Organizational Psychology Commons

Recommended Citation
O’Toole, Michael F., "The Validation of a Test Battery and Biographical Data Used to Select Key-Punch Operators" (1976). Master's Theses. 2381.
https://scholarworks.wmich.edu/masters_theses/2381

This Masters Thesis-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Master's Theses by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
THE VALIDATION OF A TEST BATTERY AND BIOGRAPHICAL DATA USED TO SELECT KEY-PUNCH OPERATORS

by

Michael F. O'Toole

A Thesis Submitted to the Faculty of The Graduate College in partial fulfillment of the Degree of Master of Arts

Western Michigan University Kalamazoo, Michigan December 1976

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
ACKNOWLEDGEMENTS

The investigator wishes to thank Dr. Bradley Huitema and Dr. Richard Schmidt for their advice and assistance in preparing this study. Sincere appreciation is extended to Dr. Frank Patzinger, without whose backing and guidance this study could not have been completed.

A special note of gratitude is extended to Mr. Allen Peat, without whose cooperation the study would never have been undertaken.

Michael O'Toole
INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in “sectioning” the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from “photographs” if essential to the understanding of the dissertation. Silver prints of “photographs” may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

University Microfilms International
300 North Zeeb Road
Ann Arbor, Michigan 48106 USA
St. John's Road, Tyler's Green
High Wycombe, Bucks, England HP10 8HR

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
MASTERS THESIS 13-9407

O'TOOLE, Michael F.
THE VALIDATION OF A TEST BATTERY AND BIOGRAPHICAL DATA USED TO SELECT KEY-PUNCH OPERATORS.

Western Michigan University, M.A., 1976
Psychology, industrial

Xerox University Microfilms, Ann Arbor, Michigan 48106
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .................................. ii

TABLE OF CONTENTS ................................ iii

INTRODUCTION. ..................................... 1

FLOW CHART OF HIRING SEQUENCE ................. 4

METHOD. ........................................... 10

PROCEDURES USED TO CALCULATE THE PRODUCTION DATA FOR KEY-PUNCH OPERATORS .... 11

RESULTS .......................................... 14

INTERCORRELATIONS OF PREDICTORS AND CRITERION . 15

DISCUSSION. ...................................... 17

APPENDIX ........................................ .26&27

REFERENCES ..................................... .28&29
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................ ii
TABLE OF CONTENTS ..................................... iii
INTRODUCTION ............................................. 1
METHOD .................................................. 10
RESULTS ................................................... 14
DISCUSSION .............................................. 17
APPENDIX ........................................... 26 & 27
REFERENCES ............................................ 28 & 29
High speed data processing through the use of computers has become a requirement for many large businesses. With this high demand for various types of data to be processed faster and more economically, a new kind of service has developed, that of "free lance" data processing. These firms do the processing for various clients that have too much processing for their own systems. But even more so, they do processing for companies that either can't afford a computer or don't have enough total processing needs to warrant the purchase of a computer. Processing of payroll records, county tax assessments and billing services are just a few of the accounts that these firms handle. Upon receiving the raw data these firms process the data into any number of output forms that the client may request, such as cards, magnetic tape, or in printed form off a high speed line printer.

Critical to this total operation are those functions performed by the key-punch operators. Any combination of low production, high turnover, or high error rate could skyrocket the costs of processing the data plus destroy the profits of the company. It is, therefore, very important to select good key-punch operators and to keep turnover to a minimum.

It was found that there were two characteristic errors being made by key-punch operators (Kirchner, 1966). The
errors fell into either perceptual (misreading) or spatial (mispunch) errors. In an earlier study (Kirchner & Banas, 1961) it was found that experienced operators were more likely to make perceptual errors rather than spatial errors. Whereas, inexperienced operators were more likely to make the spatial errors over the perceptual errors. They also found that the error rate for both types of errors was less than 7%. The company being used in this study did not have data pertaining to the types of errors being made (ie, spatial or perceptual), but did have data which showed that the total error rate was less than 3%. This was not considered to be a major concern, by either the company or by this investigator as their criterion level was 3% error rate. It should be noted though that in the case of this company each job was verified, that is, once a job was completed by one operator, it was repunched by another operator to check for errors.

The turnover of key-punch operators is a problem found in almost every data processing operation, as is typical of most clerical jobs. Lawshe and Balma (1966) report that turnover rates of 35% or higher are not unusual in these job situations. The company involved in this study has not managed to escape from this problem. In 1974 a 68% turnover rate was reported, while in 1975 they incurred a 54% turnover of key-punch operators. This company currently employs 28 full time key-punch operators.
One continuous task, therefore, is the recruiting of good key-punch operators. New people, many inexperienced, must be hired to replace those who have left for various reasons. The selection of potentially good operators from such a heterogenous pool of workers has become highly important.

Applicants for key-punch positions at this company all follow a regular selection process, as seen in Figure I. After an application blank has been completed, the prospective employee is required to take two employment tests. If the applicant has had previous key-punch experience, a work sample test is administered, if the applicant has had no previous experience, then a typing test is administered. Both of these tests are company developed and have no score attached to them, the applicant either passes or fails. At this point each applicant is briefly interviewed, as a final screening. Upon being hired, those people with previous experience are placed on the regular incentive payroll system after a two week adjustment period. Those hired with no previous experience are classified as trainees and are not placed on the incentive system for up to six months.

A major screening device used by many employment offices is some type clerical test. Bennett and Cruickshank (1949) state that while clerical positions vary considerably, many clerical jobs have a certain number of tasks in common. Although the specific tasks vary, the
FIGURE 1
Flow chart of hiring sequence

Application Blank → Tests

Experience → Tests

No Experience → Typing Test

Key-Punch Test

Fail → Not Hired

Typing Test

Interview

Not Hired

Successful

Hired → Job Performance

Interview

Not Successful

Hired → Job Performance

Successful
aptitude required for the tasks is similar.

A device used to check the clerical aptitude of applicants is in the Minnesota Clerical Test. The Minnesota Clerical Test is a test of speed and accuracy in performing tasks related to clerical work. It has been found useful for selecting clerical employees and for advising persons who wish to seek training in the clerical field.

The test consists of two parts: Number Checking and Name Checking. In each part there are two hundred items consisting of one hundred identical pairs and one hundred dissimilar pairs. The numbers in Number Checking contain from seven through seventeen letters. Separate time limits are used for the two parts, with a total testing time of 15 minutes.

The Minnesota Clerical Test was designed and standardized for adult use, the adult group included females of 17 and above and males aged 19 and above. It was originally assumed that the test was equally applicable to both males and females, but data for sex norms were subsequently compiled (Super, 1961). Andrew and Patterson (1934) also found the necessity to compile separate norms for experienced and non-experienced workers.

The Minnesota Clerical manual reports test-retest reliabilities of between .61 and .76 for clerical workers and up to .93 for students in a commercial training course. Because the Minnesota Test is a speeded test, the determin-
lation of reliability by correlating scores on odd and even items would be inappropriate (Manual).

Several validity studies have been done to determine success of key-punch operators. Using an "errorless production" score Super (1962) found validities of .32 for Numbers and .29 for Names; these were the most valid test scores in the battery correlated with the criterion. Hay (1943) found validities of .51 and .47 respectively for coding clerks and hand transcribers. The discrepancy may be the differences that lie in the criteria used in each study.

As an aid in assessing the applicant's general intelligence level, the company administers the Wonderlic Personnel Test. The Wonderlic Personnel Test, originally developed in 1942, is a fifty item, timed, spiral omnibus measure of intelligence. The Wonderlic is basically a twelve minute adaptation of the Otis Self-Administering Test of Mental Ability and was designed for employment office use as a test of general learning ability. The Wonderlic has been validated in a number of industrial situations. The original forms (A, B, C, and F) were established by using more than 30,000 adults ranging in age from 20 to 65. The norms were based upon adult groups in industry making the test particularly suitable for employment situations. Comparative norms for large groups of adults are found in tables in the manual, (Wonderlic, 1966), and are categorized according to age, sex and level of education.
Minimum occupational scores for 75 professional, office, plant staff and line personnel, and various other positions are also provided.

Wonderlic forms, I, II, IV V and EM were developed in 1959 and were based on research conducted during 1950 through 1959. In 1961, E. F. Wonderlic (1966) analyzed data from more than 625 organizations in business, industry and government. He classified data by age, sex, education and equivalent test form. He found that the age corrections provided in the 1940 edition of the Wonderlic were still valid. At this time he found sex differences and recommended separate norms. Also, educational level differences were found which then necessitated separate norms. Reportedly the Wonderlic is one of the most widely used tests in business and industry. (Lawshe and Balma, 1966).

The Wonderlic manual reports test-retest reliabilities of .82 to .94. Comparisons of odd numbered items against even numbered items resulted in correlations ranging from .88 to .94. The test-retest reliabilities are probably of more use, since even-odd comparisons are not very applicable to speeded tests (Cronbach, 1966). Droege (1970) reported that there is not nearly enough documented validation data to justify the current wide use of the Wonderlic Personnel Test in screening applicants for employment or candidates for promotion. He feels that aptitude and intelligence tests are coming under attack because of the
possibility of unfair discrimination against minority groups in the application of the test results in selection. The Wonderlic is particularly vulnerable because of its known high relationship to formal education, the use of a priori norms and the lack of documented evidence that it is a useful instrument in predicting success in many jobs. Unless the user first validated the test on a sample of his applicants or workers, he may find that he is in violation of Title VII of the Civil Rights Act of 1964.

The company being used in this study, wishing to comply with the present laws, sought to validate the selection procedure used for their key-punch operators. This study was undertaken to determine the extent to which the Wonderlic Personnel Test, the Minnesota Clerical Test and biographical data can empirically differentiate between successful and unsuccessful key-punch operators.

The original goal of this study included the plan to study the predictive efficiency of the job sample tests, but the hypotheses tested for the key-punch operators were: 1) job performance, can be predicted from test scores and biographical data obtained at the time of the final screening interview, and 2) job tenure, as defined by months on the job, can be predicted from test scores and biographical data obtained at the time of the final screening interview.

The tests used in this study were divided into two classifications: 1) test of general learning ability
(Wonderlic Personnel Test) and 2) test of perceptual speed and accuracy (Minnesota Clerical Test).
METHOD

The criteria of job success were selected for use in the validation of the selection procedure used by the company in this study. These measures were selected on the basis of the following reasons: practical significance of the criteria, accessibility of the data and the probable reliability of the personnel records.

The first measure selected was the percentage of production of base-rate for each operator. This measure was devised by taking a sampling of each operator's gross paycheck, excluding any overtime pay. This figure was then divided by the total hours that the operator worked that particular week, the result was the average hourly rate for the week. The average hourly rate is divided by that operator's base hourly rate and then multiplied by 100, which produces the percentage of base rate for that week. An example of this procedure is worked in Table 1.

Barnes (1968) presents similar concepts when discussing the relationship of "normal pace" to "average incentive pace" when 100% equals normal pace in time study procedures.

The average incentive pace was calculated for a total of twelve randomly selected work weeks during 1975. If a particular operator did not work all of 1975, than a twelve week random sampling was taken for the time that the operator was employed.
TABLE I

Procedure used to calculate the Production data for key-punch operators

\[ \begin{align*}
$110.00 &= \text{week's gross pay} \\
37 \text{ hours} &= \text{total hours worked excluding overtime} \\
$2.50 &= \text{base hourly rate} \\
\frac{$110.00}{37 \text{ hours}} &= $2.97 \text{ average hourly rate for one week} \\
$2.97 \div $2.50 &= $1.18 = \text{raw output total} \\
1.18 \times 100 &= 118\% \\
13\% &= \text{percent of production over base rate}
\end{align*} \]
The output of the key-punch operators is measured in terms of strokes. A stroke is recorded each time the key-punch operator depresses any key on the keyboard. A normal stroke rate is determined for each job or account the operators work on. This normal stroke rate is then multiplied by a difficulty factor which minimizes any discrepancies due to the difficulty of one job over another. This procedure is done by the company's Honeywell 20 series computer.

Wonderlic and Minnesota Clerical test scores, production and tenure along with the biographical data were entered onto a disk file in the PDP-10 computer at Western Michigan University Computer Center. This computer also provided the printouts of statistical analyses and other pertinent information.

For the production data, point-biserial correlation coefficients and single and multiple correlation coefficients were computed for: 1) all predictors and 2) individual and selected predictors. The point-biserial was used to force the distribution into a dichotomy of above average and below average workers.

The second criterion measure used in this study was tenure with the company being investigated. Those key-punch operators hired between January 1971 and December 1975 were included in the tenure portion of the study. Tenure is defined as the cumulative number of months the
employee had been working for the company at the time of termination or on December 31, 1975 if still employed. An employee also had to be on the job at least six months to be included in the tenure group.

Test scores and tenure and biographical data for the key-punch operators were then placed on a disk file in the PDP-10 computer, for analysis.

Point biserial and single and multiple correlations between test scores and biographical data, and tenure data were computed.
RESULTS

The first hypothesis tested for the twenty-seven key-punch operators was: job performance, can be predicted from test scores and biographical data obtained at the time of the final screening interview.

The intercorrelation matrix between production and all the predictor variables is presented in Table 2. One significant correlation is between age of the employee and the production output. There are also significant correlations between Minnesota Clerical Names and Number checking and Wonderlic test scores. The above significant correlations are significant at .05 level of confidence. None of the point-biserial correlations between production and the predictors are significant. The dichotomy for the point-biserial correlations was produced by using the average production for all key-punch operators. The multiple correlation coefficient for all the predictors and production equalled .56, this is not a significant multiple correlation at the .05 level of confidence. The test battery (Wonderlic, & Minnesota Clerical) has a multiple correlation of .31 with production and is not significant at the .05 level of confidence.

The second hypothesis tested for the key-punch operators was: job tenure, can be predicted from test scores and biographical data obtained at the time of the final
### TABLE 2

Inter-correlations of predictors and criterion

<table>
<thead>
<tr>
<th>PREDICTORS</th>
<th>X1</th>
<th>X2</th>
<th>X3</th>
<th>X4</th>
<th>X5</th>
<th>X6</th>
<th>Y1</th>
<th>Y2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.128</td>
<td></td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.263</td>
<td></td>
</tr>
<tr>
<td>Key-Punch Experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.03</td>
<td></td>
</tr>
<tr>
<td>Wonderlic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.03</td>
<td></td>
</tr>
<tr>
<td>Minnesota Clerical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.191</td>
<td></td>
</tr>
<tr>
<td>Number Checking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.218</td>
<td></td>
</tr>
<tr>
<td>Name Checking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.162</td>
<td></td>
</tr>
<tr>
<td>Criterion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Production</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.525*</td>
<td></td>
</tr>
<tr>
<td>Tenure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.069</td>
<td></td>
</tr>
</tbody>
</table>

*Significant at .05 level
screening interview.

The intercorrelation matrix between tenure and all the predictors can be found in Table 2. There are no significant correlations between tenure and the predictors. There is a significant negative point-biserial correlation between job tenure and the Wonderlic test scores of -.41 at the .05 level of confidence. The average tenure for all the operators was used to determine the dichotomy for the point-biserial correlations. The multiple correlation coefficient for all the predictors and job tenure equals .41, this is not a significant multiple correlation coefficient, at alpha equal to .05. The test battery has a nonsignificant multiple correlation, with job tenure, of .37 at the .05 level of confidence.
DISCUSSION

The current research was conducted in an attempt to develop a set of predictors what would aid in the selection of potentially successful key-punch operators at the company being used in this investigation.

With the production data criterion, a multiple correlation of .56, obtained by using all the predictors, was not significant, at .05 level of confidence. Neither was the multiple correlation of .31 between the test battery and production. It can be concluded therefore, on the basis of the data collected, the prediction of production, by use of the test battery and biographical data for key-punch applicatns does not seem to be practical, without further investigation. It was decided, then, to reject the first hypothesis.

The second criterion, job tenure, produced a multiple correlation of .41 with all predictors, which was not significant at the .05 level of confidence. The multiple correlation with the test battery also produced a nonsignificant coefficient of .37 at the .05 level of confidence. It was, therefore, also decided to reject the second hypothesis that these predictors in combination would not be useful, without further investigation. When the recorded correlations between separate predictors and criterion
were examined, the correlation between age and performance was .53 and was significant at the .05 level of confidence. A partial correlation run on the production data and age of an employee, partialling out experience, resulted in a significant correlation of .57 at alpha equal to .05. This reduced the possibility that the previous significant correlation was contaminated by experience. Shott, Albright and Glennon (1964) found that the age of office workers, when compared to their tenure with previous employers lent credence to their results, which showed high performance standards being made by older employees. The authors alluded to the possibility that these older employees may be motivated by external factors, not directly connected with their jobs. One of these factors was dependence on the job resulting from financial obligations, such as, house payments. They felt that because of some of these external factors, the older workers were more likely to produce more than their younger counterparts. Guion (1965) hypothesized that one's credit standing could serve as a motivating force on the job, particularly where pay reflects effort. In other words he is saying that the amount of money one owes could be a factor on their output.

The negative point-biserial correlation between job tenure and Wonderlic test scores was -.41. This was a significant correlation at the .05 level of confidence. Numerous authors (Guion 1965, Kreidt and Gadel 1953, and
Tiffin (1947) presented various negative relationships between tenure and a measure of general mental abilities. It was found by Kreidt and Gadel (1953) that those employees who scored high on the General Abilities Test and had routine clerical jobs tended to leave their jobs between three and twelve months (short tenure). They found that those who scored in the middle to low range and had routine clerical jobs tended to stay on the job longer than twelve months (long tenure). The company used in this present study has tended to use the Wonderlic test and select those applicants with above average scores, which could in part explain the high negative correlation.

Guion (1965) and Tiffin (1947, 1965) and Schuh (1967) cited several inherent problems in validation studies. That of present-employee (concurrent) validation, homogeneity of the sample and sample size all having an effect upon the correlation coefficients.

The present-employee or concurrent validation method, uses people who have been with the company for some time. The test battery is administered to those people whose job behavior can already be rated as "good" or "poor" according to an established criterion. Guion (1965) feels that this method is valid to the extent that it correlates with criterion measures available at the time the test is given. Guion also states that the concurrent validation method is clearly a violation of scientific principles. He argues
that present employees constitute a decidedly different population from that composed of job applicants. With ability tests, as opposed to interest inventories, the difference may be important, but concurrent validities may occasionally be useful only as "first guesses" about predictive validities. The uncertainty with the concurrent method means that no great faith should be placed on concurrent validities until more adequate follow-up studies are completed.

Tiffin (1947) states that caution should be employed when interpreting the results of a present-employee validation study. If one is not careful, he might attribute high aptitude for a particular job to the fact that an employee scored high on an aptitude test, when actually because of experience or maturation on the job, the test is actually measuring achievement. An unpleasant truism is that often times we seldom know precisely what it is that produces the validity. The concurrent validity may be higher or lower, and the difference may be large or small. There is no necessary relationship between concurrent and predictive validities; Guion (1965) feels that a tester is simply deluding himself if he uses one as an estimate of the other.

Another inherent problem with this study is the homogeneity of the sample. Because any obtained correlation coefficient depends in part on the extent of individual
differences in the group measured (Guion 1965). The size of the total variance does not influence the size of the error variance since error variance is a function of the repetition of measurement for each individual. Because reliability coefficients are smaller as the ratio of error to total variance gets larger, obtained reliability coefficients are smaller where the total variance is smaller.

Due to the nature of this study and the company's selection procedure, the employees used tended to be highly pre-selected, which becomes important in interpreting reliability data obtained from such a group. Because everybody's test score tended to be just like everybody else's (caused by the previously mentioned high cut-off for the Wonderlic scores), there was a lack of variability, which as mentioned could have an effect upon the correlation coefficients.

A final problem in interpreting the results of this study is the sample size. Because the sample size was twenty-seven, it is probable that the obtained correlation coefficients were unreliable. Guion (1965) feels that increasing the number of subjects does not have a necessarily systematic effect on the size of the obtained coefficients, but it does have a major effect upon its dependability. Schuh (1967) showed the effect of instability of correlation coefficients due to sample size. He correlated 31 different variables with tenure over four separate years.
The sample size varied no more than six subjects from year to year with a maximum of 36 subjects for any one year, for any particular correlation group. The subjects were different from year to year. He found correlations which varied from year to year, up to a $/0.34/$ difference for any given two years. Schuh showed that the combined correlations for all years was a much more stable measure than any of the correlations for one year.

It has been reported that those employees in clerical positions which gave them some freedom and responsibility in planning their own work tended to be long tenure employees. These findings are supported by Pond and Bills (1933) who reported that those employees on the easiest jobs with the highest test scores on a general mental abilities test had the highest turnover rate.

Although the company used in this study has supervisors on each shift who are responsible for the activities of the key-punch operators, management should explore the possibility of involving the key-punch operators in the scheduling and planning of their work. It is probably not feasible to undertake a total "job enrichment" program because of the nature of key-punch jobs, but minor changes that would create a more challenging and stimulating work environment should be investigated. In addition to those changes already mentioned, this company might investigate the possibility of involving the workers in changing the work envi-
ronment, wage and incentive systems, and developing a formal grievance procedure.

Another aid to help reduce the turnover rate at the company in this study would be to use an appropriate questionnaire or exit interview at the time an employee terminates. At such time that an employee terminates, she would be asked to either fill out a questionnaire or have an interview, responding to pertinent questions regarding her termination. This would not be a very reliable technique, but it might give the company officials a better insight into their turnover problem. By attending to the responses of terminated key-punch operators, it is possible that management may be able to constructively alleviate certain types of grievances which result in terminations.

There have been some recent studies done to try and validate tests for the selection of key-punch operators. Those studies have run into some critical problems, such as one reported by Shub and Campbell (1972) in which they obtained a multiple correlation of .40 with supervisory ratings. When these results were cross-validated they reported that there was virtually no correlation. In this as well as other studies this investigator has reviewed there is the problem of obtaining meaningful and relevant criterion of job success. Guion (1965) states that the solution to the problem is simple: do not weight, do not combine, use more than one concept of how performance on a
job can be evaluated. The single composite, based upon a concept of ultimate worth, is often valuable and perhaps even indispensable in some studies. More often, though, it is probably not justified. The idea of a unitary construct justifying the combination of such measures as productivity, accident rates, absences and errors into a meaningful composite has been flatly contradicted by data reported by Seashore, Indik and Georopoulos (1960).

The firm used in this study should take steps to evaluate its criterion measures by answering the following questions: there has to be a decision made as to the importance of the concept being measured, is it crucial to the functioning of a person on this job? Second, if the concept is important and an operational definition is developed, the relevance of the measure must be judged in terms of the degree to which it is related to the concept (Guion 1965).

Because of the size of the company used in this study, it is not likely they would have more than 35 key-punch operators working at any one time. If management is to make valid selections of future key-punch operators they must make the time and money investment to do a predictive validity study. When they conduct the predictive study, management might consider using a technique similar to that of Schuh's, to enable them to obtain a more reliable sample size.
In addition to answering the questions about the criterion, management should probably look at their present selection instruments. The Wonderlic test could probably be dropped as it adds little in selecting potentially good key-punch operators and may actually have a reverse effect on tenure. Management elected to choose those people who tended to score above the average for key-punch operators, which apparently resulted in a high turn-over for those very people. The Minnesota Clerical test should be used in further validity studies and selection procedures for employees. Although the correlations for this test were not significant, they did show a positive relationship with the criterion. This test measures those concepts which are important and relevant to the everyday functions of a key-punch operator. One last thing management should do to improve their present selection instruments would be to develop some objective measurement for their company-developed tests. These tests are actual work samples of on-the-job behaviors, which could give much more meaningful feedback as to the potential of an applicant than most paper and pencil tests.
<table>
<thead>
<tr>
<th></th>
<th>Above Average</th>
<th>Below Average</th>
<th>Production</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above Average</td>
<td>11%</td>
<td>33%</td>
<td></td>
</tr>
<tr>
<td>Below Average</td>
<td>33%</td>
<td>22%</td>
<td></td>
</tr>
</tbody>
</table>

APPENDIX A
Job Tenure
APPENDIX B

Wonderlic Scores

<table>
<thead>
<tr>
<th>Tenure</th>
<th>Below Average</th>
<th>Above Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above Average</td>
<td>25%</td>
<td>14%</td>
</tr>
<tr>
<td>Below Average</td>
<td>18%</td>
<td>40%</td>
</tr>
</tbody>
</table>
REFERENCES


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Schuh, A. Application Blank Items and Intelligence as Predictors of Turnover. *Personnel Psychology*, 1967, 59-64.


Stephens, J. B. "The Validation of Tests used to select secretarial and clerical employees at Western Michigan University." Thesis, Western Michigan University, 1968.


