The predictive validity of a police officer selection program

Neil Bingham Davidson
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds
Part of the Social Control, Law, Crime, and Deviance Commons, and the Social Psychology and Interaction Commons

Let us know how access to this document benefits you.

Recommended Citation
Dissertations and Theses. Paper 2276.
https://doi.org/10.15760/etd.2273

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Title: The Predictive Validity of a Police Officer Selection Program.

APPROVED BY MEMBERS OF THE THESIS COMMITTEE:

Milton Davis, Chairman

Robert Pawloski

Morris Weitman

This study was designed to determine the predictive validity of a police officer selection program and identify the contribution made by each major selection device to the total program. Police officers employed by the Portland Police Bureau who had completed three years of post-probationary employment were randomly assigned to a validation group and a cross-validation group on a two for one basis respectively. Beta weights were computed for the written test, interview and psychological scores in the validation group. The regression formula was then applied to the data in the cross-validation group. A cross-validation R of 0.12
was obtained between the predicted performance criterion scores and the actual performance scores. When the interview variable was removed from the equation the cross-validation $R$ increased to 0.16. Neither validity coefficient reaches statistical significance. Reasons were offered for believing that the low magnitude of the coefficients was attributable to restriction in range in the predictor variables and the unreliability of the criterion variable.

The spread in written scores and psychological scores for terminated police officers who are classified as occupational failures was found to be significantly greater than for officers who resign in good standing and officers who stay with the department. Police officers who later received promotions were found to have scored exceptionally high on the entry written test.
TO THE OFFICE OF GRADUATE STUDIES AND RESEARCH:

The members of the committee approve the thesis of Neil Bingham Davidson presented July 29, 1975.

Milton K. Davis, Chairman

Robert Powloski

Morris Weitman

APPROVED

Ronald E. Smith, Head, Department of Psychology

David T. Clark, Dean of Graduate Studies and Research
THE PREDICTIVE VALIDITY OF A POLICE OFFICER SELECTION PROGRAM

by

NEIL BINGHAM DAVIDSON

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE
IN
PSYCHOLOGY

Portland State University
1975
ACKNOWLEDGEMENTS

I would like to express my appreciation for the support and encouragement that has come from friends and associates. Special thanks should go to the members of my thesis committee for their ideas and suggestions. The late evening telephone calls from Dr. Milton Davis, my thesis advisor and committee chairman, served to spur me on.

I would like to thank Mr. Jon Stephens, present Secretary to the Portland Civil Service Board and Mr. George McDonald, past Secretary to the Board, for allowing me access to the official personnel records of the City of Portland. Thanks should also go to the Chief of Police and his staff for their support and assistance.

I would like to express special thanks to Lieutenant Andrew Crabtree (retired). His professionalism and his dedication to his field instilled in me a profound respect for those who choose to serve in this demanding and often unappreciated line of work. A selection system that would consistently result in the employment of people of his caliber should be the ultimate goal.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vi</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II LITERATURE REVIEW</td>
<td>4</td>
</tr>
<tr>
<td>III METHOD</td>
<td>13</td>
</tr>
<tr>
<td>IV RESULTS</td>
<td>36</td>
</tr>
<tr>
<td>V DISCUSSION</td>
<td>46</td>
</tr>
<tr>
<td>VI SUMMARY</td>
<td>53</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>56</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>57</td>
</tr>
</tbody>
</table>
### LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLES</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I  Status of 254 police officers hired from</td>
<td>29</td>
</tr>
<tr>
<td>January, 1965 to June 1970</td>
<td></td>
</tr>
<tr>
<td>II Comparability of validation and cross-validation samples</td>
<td>37</td>
</tr>
<tr>
<td>III Intercorrelation matrices</td>
<td>38</td>
</tr>
<tr>
<td>IV Beta weights and multiple R's</td>
<td>40</td>
</tr>
<tr>
<td>V Predictor variables and standard deviations for four groups of</td>
<td>41</td>
</tr>
<tr>
<td>subjects based on employment status</td>
<td></td>
</tr>
</tbody>
</table>


LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURES</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Flow chart showing sequential selection strategy utilized in hiring police officers from January, 1965 to June, 1970</td>
<td>17</td>
</tr>
<tr>
<td>2 Distribution of performance ratings when converted to standard scores within year groups</td>
<td>26</td>
</tr>
<tr>
<td>3 Scatter diagram showing relationship between predicted performance scores and true performance scores utilizing written and psychological scores only</td>
<td>43</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

The effective functioning of law enforcement organizations is of vital importance in preserving the social fabric of our society. Probably no other factor has a greater bearing on the functioning of law enforcement organizations than the quality of the people who are employed to enforce the law. According to Katzenbach, then Chairman of the President's Committee on Law Enforcement and Administration of Justice Task Force:

"It has often been said that policing a community is personal service of the highest order, requiring sterling qualities in the individual who performs it." (p.125)

The possible consequences of hiring an inept or indifferent police officer are frightening both in terms of personal loss and the forfeiture of selection training investment. Over ten years ago the average cost of training one police officer to a minimally acceptable level of proficiency was estimated to be approximately $10,000. In the years since then, the costs have no doubt spiraled to at least twice that amount.

Given the importance of identifying, attracting, and employing the most qualified candidates for the law enforcement field no one would wish to see hiring performed on a haphazard basis. In fact such hiring is usually performed
in a systematic and organized manner. Narrol surveyed 61 cities with populations greater than 150,000 according to the 1959 census in order to estimate the extent to which formal psychological and psychiatric techniques were being employed in the selection of police recruits (1963). Of the 55 cities responding, 40% used standardized IQ test, 87% used non-standardized police selection tests, 16% used standardized group personality tests and 16% used a psychiatric interview. Small percentages of the responding cities used other instruments. All of the cities responding used at least one test in selecting candidates for employment. However, in spite of the wide use of formal selection programs, very little effort has been made to evaluate the instruments and selection programs for effectiveness.

Efforts made to establish the validity of hypothesized predictors have varied from occupation to occupation. Generally speaking, by far the greatest effort has been made in the area of production work using as subjects those people actually involved in production and as a criterion measure, the varying aspects of work output. Non-production studies predominately deal with selection for sales positions, which of course, also afford a clearly defined criterion measure. Service oriented occupations have received relatively little attention as far as test validation is concerned. Undoubtedly this uneven emphasis has been due to the difficulty involved in defining criteria, reliably measuring criteria.
once they are defined, and finally obtaining a consensus as to the relative importance of various criteria for any given job.

Authorities in the field of tests and measurements continually emphasize that selection instruments must be validated periodically in each setting in which they are administered. Research indicates that the validity of a given instrument may be affected by such factors as: the locale in which it is used, the changing make-up of the population from which the potential employees are drawn, and the changes that occur with time in the nature of the work to be performed by individuals in a particular occupation.

Reflecting the difficulty encountered by the police in predicting the suitability of candidates who seek to enter the police field, the studies of police selection have been sporadic and the resulting conclusions have had to be qualified. The following chapter is a review of the few major studies which constitute the major research to date.
CHAPTER II

LITERATURE REVIEW

An early summary study was published in 1949 (Ghiselli). For the purposes of the study police, firefighters, guards, and other related workers were combined into the General Occupational Classification (GOC) protective service occupational group.

The median coefficient for different kinds of tests as they related to trainability were: intelligence abilities, .35; spatial and mechanical abilities, .34; perceptual accuracy, .30; and personality traits, .13. For the criterion of job proficiency Ghiselli found the following coefficients: intelligence abilities, .23; spatial and mechanical abilities, .16; perceptual accuracy, .19; and personality traits, .21.

Unfortunately the characteristics of the selection programs that had generated the data were not described, nor were the locales in which the tests were used. The criteria used, other than the fact that they were reducible to trainability and job proficiency, were not described. At what point in the subjects' careers the criterion data was gathered was not given. While not specifically mentioned it can be assumed that most of the studies were concurrent in nature. No mention was made of selection ratios and the magnitude of
restriction of range. Most of the validity averages were based on only a few correlations, with the exception of intelligence which reflected eight coefficients. The number of subjects in all of the studies combined was less than 500. However, in spite of the shortcomings described above, the study did point out for the first time that intelligence relates positively to job trainability and performance and that a number of other measures designed to measure other attributes offer potential for predicting these two important criteria.

At about the same time that Ghiselli's summary study appeared, a study was conducted on St. Louis patrolmen selection techniques (DuBois and Watson, 1950). This study is still considered by many to be the best example of validation research in the law enforcement field. DuBois and his associates administered a battery of selection and experimental tests to the two entry classes (N-72 and N-57) of probationary patrolmen at the St. Louis Police Academy. The criterion measures were four in number: final grades in the academy, scores on an achievement test covering police science taken 10 weeks after graduation from the academy, marksmanship, and supervisor ratings taken after 10 weeks on duty. The ratings were based upon an 11-trait rating scale with five descriptive steps for each trait. Only the validities against the overall service rating were reported.

The Army General Classification Test (AGCT) was found to be a good predictor of academic performance (beta .41 and
beta .28 respectively). The Police Aptitude Test, specifically designed for use in the study, was found to also predict academy grades (beta .11 and beta .28 respectively). The Police Aptitude Test, which is very similar to the written test used by the City of Portland, was found to correlate at .53 with the AGCT in a very restricted range. DuBois found no relationship of significance between the individual tests and supervisor ratings. A combination of three non-verbal tests and the Rosenzweig Test yielded a multiple correlation of .29 with the ratings, a relationship significant at the .05 level of confidence.

While basically very well designed the study does have a few shortcomings. The supervisor ratings were gathered one time only, shortly after completion of the academy and long before job performance could possibly have stabilized. The cross-validation effort was limited to the academy grades criterion.

Personal and psychological characteristics of Portland policeman and fireman applicants were described by Matarazzo, et al (1964). The applicants (policeman N-116, fireman N-127) were psychologically screened as part of the regular hiring process during the three year period extending from 1959 to 1961. It was found that the police candidates had an average IQ of 112 (WAIS Full Scale) which falls at the eightieth percentile in the general population. It is of interest that DuBois found that St. Louis police candidates on the average achieved a score of 118 on the AGCT since that is also a
score corresponding to the eightieth percentile in the general population.

On various scales designed to measure emotional adjustment Matarazzo found the applicants to be at the very healthy end of the scale. No pathologically high scores were found on the MMPI. As to personality needs the applicants were higher than average in their needs for achievement, exhibition, intraception (ability to analyze and understand the feelings of others), dominance, endurance and heterosexuality (masculine interests), and lower than average in autonomy (need to work independently), succorance (need for encouragement, kindness, and help from others), nurturance (need to forgive, sympathize with, or to help friends and strangers who are sick or in trouble), and aggression (need to criticize others, or tell them off, or get revenge). The applicants' vocational interests seemed to be oriented toward the social services, in other words, toward jobs that involve working with people.

The purpose of the research was to point out the differences and similarities between policeman and fireman applicants and show how both groups compared with the normal population for each standardized evaluation instrument. The study dealt solely with predictors and did not attempt to relate the predictors to a criterion.

Biographical information contained in the personnel files of fourteen law enforcement agencies was scrutinized by Levy to determine if bio-data could distinguish between
officers who leave for cause (Failures), officers who leave though in good standing (Non-failures), and officers who stay in the law enforcement field (Currents). (1967) In total, information was analyzed for 690 Failures, 643 Non-failures, and 1,333 Currents. All variables were used on which information was found in at least half of the personnel folders. Application of this criterion resulted in 140 usable variables.

Non-failures were youngest and most educated at the time of appointment. Current employees were oldest and least educated at the time of appointment. The occupational failures tended to have more vehicle code and other violations, a greater number of marriages, greater tendency to have been fired or asked to resign from previous positions and greater tendency to have grown up in a family from which the father was absent due to death, divorce, or emotional trauma.

Chicago Police Department beat patrolmen were the subjects of a study conducted by the Industrial Relations Center of the University of Chicago (Baehr, et al, 1986). The purpose of the study was to identify test instruments that related significantly to performance and to identify distinctive patrolmen types based on different types of field performance. The assumption underlying the second aspect of the project was that there are a variety of successful and unsuccessful patterns of field performance.

The extensive battery of tests selected for use in the study covered as wide a spread of human behavior as possible.
Baehr categorizes the instruments as follows:

1. **Motivational Measures**
   - Dimensions of objective background data
   - Dimensions of work interests indicating strength, flexibility, and vocational aspiration level of occupational interests

2. **Intellectual Measures**
   - Primary mental abilities in the area of reasoning, language facility, and visual perception.
   - Special aptitudes, such as creative potential and insight into social situations.

3. **Behavioral Measures**
   - Dimensions of the relatively permanent temperament traits or characteristic modes of responses of individuals in both normal and pressure situations
   - Dimensions of personality functioning as measured in tests purporting to represent various personality systems

The Chicago Police Department semi-annual performance ratings were adopted as one of the major criterion measures for the study. The ratings were routinely prepared by district supervisors for the administrative purposes of the department. A paired-comparison appraisal technique, also intended to serve as a global measure of patrolmen performance, was used conjointly with the department's inhouse rating system. Six other measures were included: tenure, awards, complaints, disciplinary actions, attendance and number of arrests.

The performance of 2327 men was rated by 253 supervising sergeants and field lieutenants. Of the patrolmen
rated 60 percent indicated a willingness to participate in the examination process. Volunteers who had been rated in the top third or bottom third were actually selected to participate. In two separate waves 490 patrolmen were tested (242 in February, 1967 and 248 in July, 1967).

Very high and significant multiple correlations were found between performance in the test battery and all eight criterion measures. The general conclusion was that ideal attributes for successful performance all relate to stability and lack of impulsiveness. Eight distinctive sub-groupings emerged within the sample based upon tenure and performance characteristics. Of considerable interest was the fact that race seemed to act as a moderator variable. The degree of the relationship between test scores and performance increased when subgroups of white patrolmen and black patrolmen were treated separately. Prediction formulas derived from the results of testing black officers could not be used on caucasians and relationships between the predictor and criterion variables for caucasians did not hold up when applied to blacks.

The study was concurrent in nature so applicability of the results to candidates for employment is questionable. Unfortunately the test instruments selected for use in the study were very dissimilar to those actually used in police selection. Most of the instruments had been designed by psychologists at the University of Chicago.
A recent validation study of two aptitude tests was conducted on a consortium basis in California and Nevada (Wollack, et al, 1973). In total, 83 protective service organizations in the two states participated. The two tests were the Police Officer A-1 Form prepared by the International Personnel Management Association (this is a revision of the written test used by the City of Portland) and the Law Enforcement Aptitude Form 51X prepared by the Cooperative Personnel Services, California State Personnel Board.

The performance evaluation scales used as criteria consisted of five cognitive dimensions: communication skills, problem-solving skills, learning ability, observation skills, and judgment under pressure; and three non-cognitive dimensions: dependability, desire for self-improvement, and attitude. The non-cognitive performance dimensions were included to serve as an experimental control. Supervisors independently completed performance evaluations for their employees. Following the independent ratings they then met in small groups and established a final and official consensus rating for each employee.

The sample included in the analysis phase of the project consisted of 396 white employees, 41 black employees, and 53 Spanish-surname employees. Due to the small number of female employees in the original sample it was necessary to omit them from the analysis. Tenure was controlled by restricting the sample to those employees at the entry-level
who had not less than one year nor more than six years of law enforcement experience.

The study revealed that both test instruments are significantly correlated with important elements of job success for entry-level law enforcement positions. An analysis of covariance for differential validity indicated that the test scores have a common meaning for all of the three racial groups included in the study. However, the validity coefficients were of smaller magnitude for the white employees relative to the validities for the two minority samples.

The shortcomings of this study are those associated with concurrent design. The supervisors at the time that they made their ratings undoubtedly knew that the information would be matched with test results. The study only dealt with one selection device, the written test. However, a written test is usually just one of a number of selection steps in the hiring process.
Methodological Considerations

Criterion related validity studies are traditionally classified into two general types: predictive and concurrent. In predictive studies the predictor variable data is gathered prior to hiring and the performance criterion data is gathered at a designated time following employment. Ideally the predictor instruments are administered to all applicants and then all applicants are hired without regard to performance on the selection instruments. Since applicants usually far outnumber the positions to be filled this optimum strategy is rarely feasible. In predictive studies it is far more often the case that the actual instruments to be studied must be utilized in selecting the subjects or other instruments similar to those under study, are administered. Concurrent studies are distinguished by the fact that the instruments to be studied are administered after employment at about the same time that the criterion data is gathered.

Predictive studies are considered to be much more desirable than concurrent studies since all tests must ultimately be predictive. Evidence for concurrent validity does not constitute evidence for predictive validity. When
significant relationships are found between predictor and
criterion data gathered at the same point in time the find-
ings do not generalize to the situation where years may
separate the gathering of predictor data from the gathering
of criterion data. Present employees, knowing that employ-
ment does not hinge on their performance on the experimental
instruments, do not approach the task with the same mental
set as applicants who are actually vying for employment.
Present employees often resent taking tests and may inten-
tionally do poorly thus adding further distortion to the
results.

It is particularly with respect to the rating process
that the predictive form of criterion related validity
truly is superior to concurrent validity. In predictive
studies rating supervisors typically do not have access to
data generated by the selection process. Since criterion
data is usually generated as just one of many routine admin-
istrative activities the rating supervisor is not sensitized
to the fact that the ratings will be used in selection
instrument research. In concurrent validation raters are
usually aware, in fact they are often purposely informed,
that their ratings are being solicited for the express pur-
pose of checking relationships between ratings and perfor-
manace on experimental tests. In other words clear independ-
dence does not exist between the criterion variable and the
predictor variable. In concurrent studies the accusation
can always be made that raters are merely reflecting in their ratings their expectations about how various subjects will fare in the testing process.

Cross-validation is a rarely utilized but exceedingly important methodological practice. The technique calls for the replication of findings which could have occurred by chance or might have been due to sampling errors or other errors of measurement. It shows how results obtained in one group may be expected to apply to other similar groups.

In practice the total sample to be studied is randomly divided into two sub-samples. One sample is designated as the cross-validation group (hold-out group). One expert in the field of validation recommends that the validation group be somewhat larger (perhaps two-thirds of the total sample) in order to assure greater statistical stability due to the larger sample size (Dunnette, 1966).

Cross-validation is particularly important in interpreting validation results when multi-regression analysis is to be utilized in the study. The coefficient of multiple correlation (multiple R) derived by a regression equation indicates the strength of the relationship between the observed criterion scores and a prediction of these scores based on a proper weighting of the selection variables. Beta weights and the resulting multiple R tend to be greatly affected by minor chance variations in any or all of the correlation coefficients. Multiple R's tend to be inflated values (Brown, 1970). This tendency spuriously increases the estimated
accuracy of predictions over that which can actually be realized when the findings are applied to new observations. Estimates of the validity of a battery of tests weighted by multi-variant techniques should be applied to a second comparable group. This process provides an uninflated estimate of the validity of the test battery.

Since cross-validation essentially removes the error in the multiple R that is attributable to working with optimum weight the resulting coefficient will generally be lower than the original coefficient. However, the cross-validation coefficient is a better estimate of the true degree of the relationship between the composite of prediction variables and the criterion variable.

**Predictor Variables**

The successive hurdles selection program utilized by the City of Portland in the employment of Police Officers can be illustrated by a flow chart (see Figure 1). Candidates wishing to be considered for employment start the process by submitting applications to the Civil Service Office. Applications are then checked for compliance with age, education and felony conviction standards. (See Appendix A for basic standards in effect over the period that subjects in this study were hired.) Conceptually a successive hurdles program involves application of tests one at a time and elimination of applicants with each test who do not score at a satisfactory level. During the period of time in question
FIGURE 1

FLOW CHART SHOWING SEQUENTIAL SELECTION STRATEGY UTILIZED IN HIRING POLICE OFFICERS FROM JANUARY 1965 TO JUNE 1970.

Step 1 Applications received by Civil Service

Applicants take written test

Step 2 Applicants take written test

N = 2,597 Passed 1,464
X = 87.2 Failed 1,133
SD = 14.7 Passing Point 84

Pass Fail (terminate)

Step 3 Applicants take physical agility test

Pass Fail (terminate)

Step 4 Applicants take oral interview

N = 968 Passed 549
X = 45.9 Failed 419
SD = 10.0 Passing Point 44

Pass Fail (terminate)

Step 5 Applicants take medical examination

Pass Fail (terminate)

Step 6 Applicants take psychological examination

N = 426 Passed 250
X = 82.2 Failed 176
SD = 5.0 Passing Point 80

Pass Fail (terminate)

Step 7 Applicants appointed to Department

N = 254
the program was designed to reflect logical and rational assumptions about the police officers' tasks and responsibilities rather than on statistically demonstrated validity. The order of presentation of instruments in the sequential selection strategy was also governed by practical and economic considerations.

All candidates who submit acceptable applications are administered a commercial police aptitude test published by the Public Personnel Association (now International Personnel Management Association) entitled: Policeman Test (10-A). The one hundred twenty multiple choice question instrument is designed to measure basic cognitive attributes such as general information mastery, practical judgment, memory span and reading comprehension (see Appendix M for analysis of police officer test by question type). Since knowledge of the law and specialized techniques in enforcing it is generally known only by persons with police experience, the material in the test is quite general. The questions are couched in law enforcement terminology and crime situations are utilized to give the instrument a touch of verisimilitude. The a priori assumption of the test is that the higher the score the greater the probability that the candidate will be able to learn the various intricacies of the job given relevant and sufficient training. Therefore, while the test is oriented towards the cognitive skills, and correlates at a high level with standardized IQ tests, it is not a pure intelligence test, nor is it a job knowledge test. The test
is timed for administrative convenience (90 minutes), but since most candidates finish well within the time provided, it can be considered a power test rather than a speed test. The passing point during the period of time dealt with in this study was approximately the mean score (see Appendix H). An odd-even split-half estimate of reliability within the range of subjects hired is 0.74.

Following the successful completion of the written test, candidates are required to take a physical agility demonstration test (pass-fail only) consisting of three basic exercises designed to measure the general level of body conditioning, strength, and endurance. Specifically, the three events used during the period dealt with in this study, in the order of administration, were eight chin-ups (arm and upper torso strength), 32 sit-ups (mid-torso strength), and a quarter mile shuttle run (lower torso strength and general state of conditioning). Failure in any of the three events was disqualifying. Only a small fraction of candidates participating at this stage in the selection process fail to demonstrate the required proficiency.

The next phase in the selection program is the oral interview. The interviews are conducted by a three-member board typically consisting of a police officer from the Personnel Division of the Portland Police Bureau and two personnel analysts from the staff of the Civil Service Board. The interview sessions are scheduled to last forty-five
minutes per interviewee and rarely deviate significantly from that length. The interview coverage is intended to be broad in scope and intensive in depth, delving into such areas as: work history, family background, education, reasons for desiring a career in law enforcement, attitudes toward racial minority groups, long term vocational expectations, etc. The ratings given, reflect the interviewer's personal evaluation of the information transmitted by each respective candidate. The scores on the written test are not available to the interview board nor are they available to the applicant until several days after completion of the interview since possession of such knowledge before or during the interview could conceivably introduce an error bias into the proceedings. The interviews are patterned and structured to the extent that each interviewer asks essentially the same questions in each interview. Therefore, each interviewee experiences essentially the same treatment. Each rater individually marks each candidate on three configurations of attributes: appearance, manner and bearing; comprehension and presentation of ideas; and attitudes toward the position. A five point graphic rating scale, ranging from deficient to superior is provided for each of the three clusters of attributes. The raters are also required to rate each candidate on a global impression of overall capability. This global rating, also on a five point scale, receives twice the weight assigned to each of the other three categories. The combined possible raw score is 75. A raw score of 44 is the
minimum passing score. If all ratings for a candidate were "satisfactory" (average) the total raw score would be 45. An estimate of the intra-rater interview reliability within the restricted range of employed subjects is 0.60. (See Appendix C for a sample of the interview rating sheet.)

The candidates who are to undergo further screening after the interview, are taken from the top of the civil service eligible list. The ordering of individuals on the list is determined by ranking the composite scores of those who pass all phases up to and including the interview. The converted written score and the converted interview score are weighted at 60% and 40% respectively and are combined to establish the candidate's official score. The weights reflect subjective estimates of relative value. At the time the composite score is established, five points are added as veteran's preference credit for those candidates who qualify.

Following the placement on the graded list of eligibles, candidates are scheduled for a medical examination and a Police Bureau conducted background investigation. Both processes are extremely thorough but only result in a small number of candidate rejections.

Candidates who are approved in the background check and medical examination are next scheduled to undergo psychological assessment. The psychological phase consists of the administration of a battery of tests and inventories followed by a clinical interview. The interview lasts from forty-five minutes to an hour per candidate. The instruments
included in the battery are changed from time to time depending on the research interests of the psychologist and his colleagues. At the time that the last subjects included in this study were assessed, the battery consisted of the following instruments: the Wechsler Adult Intelligence Scale, the Rorschach Inkblot Test, the Rotter Sentence Completion Test, the Saslow Psychosomatic Inventory, the Adorno Authoritarian P-Scale, the Strong Vocational Interest Blank, the Minnesota Multiphasic Personality Inventory, and the Halstead Reitan Neuro-Psychological Battery. During the clinical interview the psychologist seeks to obtain salient information that would provide insight into the candidate's psychological makeup and the possible presence of undesirable emotional predispositions. Each candidate is assigned a score based on the psychologist's interpretation and synthesis of the various test responses and the information garnered from the interview. The evaluation and weighing of test scores and the value placed on the information from the interview is a clinical judgment on the part of the psychologist. Those candidates who receive an overall score of 80 points or higher based on a hundred point scale are recommended for hiring. Those candidates scoring below the passing point, with rare exception, are eliminated from further consideration.

Practically all candidates who pass the psychological screening are employed. Typically, a few completely processed candidates request removal from the list prior to
appointment due to other job offers or a change in attitude about the desirability of a career in law enforcement.

Although rarely recognized as such, a final phase of the examination actually comes after appointment. During the first twelve months (recently changed to eighteen months) the rookie officer is on probation subject to termination without recourse to a dismissal appeal. During the probationary period all officers attend a three month law enforcement academy. The subject matter covered at the academy includes: criminal law, juvenile procedures, crime investigation, offensive-defensive tactics, patrol techniques, traffic operations, firearm procedures and community-police relations. During the non-academy phase of the probationary period, the rookie officer is rotated through various precincts and work shifts. During this period the probationary officer is under the continuous supervision of an officer-coach. Due to the complexity of the work, a police officer is not considered to be a producing member of the Bureau until the probationary period is successfully completed and a permanent work assignment is made.

The Criteria

The measure of actual performance thought to be most appropriate for use in this study is the composite of six post-probationary bi-annual supervisory evaluations covering the initial three year period of employment immediately following the probationary year. The Portland Police Bureau
performance evaluation report form contains graphic rating scales for six attributes: quality of work, quantity of work, work habits, adaptability, personal relations and an optional category. Each attribute can be scored from zero to three, two being deemed as competent. (A sample of the report form and the accompanying instructions are shown in Appendix D and E respectively.) The officer being rated must see the report and have an opportunity to discuss it with the rater. The report must also be seen and signed by the division head before it is routed through channels to the police personnel office.

The optional category is utilized now and then at the discretion of the rating supervisor. Since it is rarely used, wherever it is present it will be omitted for the purposes of this study.

The raters for the subjects in this study are sergeants. Due to the nature of police organizational structure sergeants have the greatest opportunity to observe work performance of street patrol officers on an ongoing basis. Due to the highly complex, technical, non-objective and independent nature of law enforcement work, only the observations of the immediate supervisors seem to have the potential for providing accurate information about individual performance.

The performance evaluation form was designed and adopted for use by the department itself; therefore, the attributes contained on it should be relevant to the goals and mission
of the organization. The form is essentially the way the organization has chosen to summarize employee performance. Evaluation scores derived from the forms should provide a relevant index of job success.

The criterion score for each subject reflects performance over a three year period of employment. The composite score is made up of six separate ratings taken at six month intervals. By controlling the amount of experience across the candidate field, potential error due to measuring performance periods of varying lengths of time is avoided. Reports summarizing short and long periods of time may give very different impressions of the same people. It should be pointed out that the rating supervisors had no way of knowing the subjects' scores on the various instruments in the selection program. At the very minimum the first performance evaluation followed the testing program by at least a year. Possible contamination of the criterion from knowledge of the testing scores was also minimized by the fact that scores were retained by the civil service office at a location physically separate from the police headquarters.

The actual range of ratings was 44 to 87. A score of 60 was the equivalent of straight 2 ratings on the six rating sheets. Only seven employees out of 205 failed to achieve a score of 60, indicating that whatever discriminative value is present it is essentially between the 2 level and the 3 level on the rating sheets. Because the obtained scores
fall within the scale range between "competent" and "superior" the ratings are actually undefined. As will be noted in Figure 2 the distribution shows positive skewness. For the purposes of this study the criterion composite scores were changed to standard scores within year groups to control for possible systematic year to year changes in supervisor rating tendencies.

In addition to employee performance, another criterion that is extremely important to an organization is job tenure. Employees who leave an organization prior to completing the standard career represent a loss of a considerable investment in training. A number of police officers hired during the five and one-half year period covered by the study.
terminated prior to June 1975. The officers who resigned voluntarily, and, because of acceptable performance, left to the sincere regret of the employer (Non-failure terminations), represent direct loss to the organization because new officers must be selected and trained to take their place. Police officers who were fired or forced to resign due to behavior intolerable to the employer, whose departure did not elicit any sorrow on the part of the employer (Failure terminations), represent a double loss to the organization because they not only have to be replaced by new officers but their own training and salary expense was not defrayed by acceptable performance while they were on the force.

Subjects

The research model to be used in this study requires the stability that can only be obtained by utilizing a large sample. Another requirement is that all subjects must have been subjected to the same hiring standards and procedures. Uniformity of treatment is critical if error is to be minimized and the results of the study are to be utilized to make personnel decisions about future samples from the candidate population.

The City of Portland police hiring program during the period extending from January, 1965 to mid-year 1970 satisfies both requirements described above. During the five and one-half year period 254 police officers were appointed to the department. By June, 1975, 205 employees in the sample
had completed the probationary period and at least three years of post-probationary employment. (In Table I a comprehensive and detailed accounting of the employment status within the total sample is presented.) To achieve 254 hires, the following activity took place at the three major processing steps in the selection program: 2,597 applicants took the written test, 968 applicants were interviewed, and 426 applicants were psychologically assessed. (For the complete derivation of the 254 police officers comprising the sample see Figure 1 and Appendices G through I.)

During the five and one-half year period the basic selection program remained essentially unchanged. Only a few minor administrative adjustments were made to the program. A major change was instituted at mid-year 1970. The education requirement was increased from a high school diploma or General Equivalency Diploma (GED) to two years of college. The Schools and Colleges Aptitude Test (SCAT) was added to the battery to serve as a waiver mechanism for candidates who had not completed the college requirement. January 1965 seemed appropriate as a cut-off for the sample to be studied since to have extended even further back would have brought into the study performance data over ten years old. Since informal organizational procedures, work standards, law enforcement philosophy and community ethos undoubtedly evolve and change over time it seemed appropriate to limit the data to that accumulated over one decade.
**TABLE I**

**STATUS OF 254 POLICE OFFICERS HIRED FROM JANUARY 1965 TO JUNE 1970**

<table>
<thead>
<tr>
<th>VALIDATION PHASE</th>
<th>TENURE PHASE</th>
<th>Medical Separation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Current</td>
<td>Non-Fails</td>
</tr>
<tr>
<td>Four years of post probationary employment completed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validation Group</td>
<td>125</td>
<td>6</td>
</tr>
<tr>
<td>Cross-Validation Group</td>
<td>63</td>
<td>2</td>
</tr>
<tr>
<td>Four years of post probationary employment not completed</td>
<td>-</td>
<td>16</td>
</tr>
<tr>
<td>sub total</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>188</td>
<td>58</td>
</tr>
</tbody>
</table>

* Three Police Officers were omitted due to incomplete data and one was omitted due to an atypical work assignment.
All subjects were males between the ages of 21 and 34 at the time of employment with a mean age of 23.9 years and a standard deviation of 2.8 years. The age distribution for the sample is very similar to that obtained for 116 police officers hired between 1959 and 1961 indicating that age characteristics of newly appointed police officers in the Portland Police Bureau are stable over long periods of time (Matarazzo, 1964). Eighty-five percent of the police officers were between the ages of 21 and 26 at the time of appointment indicating that entry age is very homogeneous across the sample. Education levels attained at the time of employment ranged from completion of high school or attainment of a GED (N-116, 45%) to graduation from college (N-9, 4%) (See Appendix N). As in the case of age, the educational distribution was remarkably similar to results obtained in Matarazzo's analysis of an earlier sample. The average educational attainment was slightly less than one year of college. Generally the subjects were reasonably homogeneous as to their age and education at the time of appointment to the Police Bureau (see Appendix O).

Female police officers appointed during the five and one-half year period were not included in the study. During the period of interest female officer employment was based upon standards and selection procedures varying significantly from those used in the employment of male officers. Female officers were selected and hired to carry out a very
specialized type of law enforcement activity that primarily involved work with female and juvenile clienteles. Female officers did not participate in standard police patrol activities. Since the employment of the last subject to be included in this study, the policy of administering separate and specialized hiring procedures and work deployment assignments for females has been discontinued.

The subjects to be included in the study were involved in similar work and were administratively treated in much the same way. Following the completion of probation the officers were assigned to uniformed patrol duties in geographical patrol districts within the city. (See Appendix F for the Police Officer job description.) After three years of post-probationary employment the comparatively similar homogeneous nature of the work begins to disappear as officers through promotion are reassigned to such specialized law enforcement activities as first level supervision (sergeant) and criminal investigation (detective). In addition to promotions, as tenure increases turnover due to resignations, injuries, and forced separations increases which further distorts the relationship of the predictors with the performance variable.

Another limiting factor that helped to define the appropriate group to be included in the sample was the probationary period. The probationary period, the first twelve months of each officer’s employment, is not representative of succeeding years on the force. During the twelve month
period each officer is primarily involved in training. By tradition, rookie police officers are rated in relationship to the performance of veteran police officers, the consequence of which is that probationary ratings are uniformly low. The range in the ratings of the probationaires does not provide adequate differentiation. In the case of the probationaires the supervisor-raters are senior police officers rather than sergeants. For these reasons data generated for the subjects during the first year of employment is appropriately excluded from the analysis.

By restricting the subjects to those who had completed three years of post-probationary employment the performance reports become broadly comparable and provide sufficient tenure to allow for informed supervisory evaluation. By limiting the sample on the basis of amount of experience, tenure which may be a factor that affects employee performance is controlled. However, the period of employment was long enough for the subjects to demonstrate their performance capabilities on the job.

Of the 254 police officers hired during the period of the study 62 terminated from the Police Bureau by June 1975. Of the terminated officers, 24 had resigned although in good standing (terminated non-failures) with the organization and 34 had left the organization not in good standing (terminated failures). The remaining 4 terminations were due to medical reasons. Since 16 of the terminated officers had completed
the three years of employment following probation, they are also included in the correlation phase of the study.

The Procedure

The police officers who were employed during the stipulated period of time and who had completed three years of post-probationary employment (N=205) were randomly assigned to a validation (developmental) group and cross-validation (hold-out) group starting with the subjects first hired on a two-for-one basis respectively. This distribution, resulting in a validation group twice the size of the cross-validation group is desirable because of the stability it lends to the validity coefficients (Dunnette, 1968). The procedure of alternating the assignment along a temporal continuum tends to control for possible tenure biases in the criterion variable.

Scatter diagrams showing the bivariate relationship between each predictor variable and the criterion were examined for evidence of curvilinearity and homoscedasticity. Means and standard deviations in the validation group and cross-validation group were checked to determine the comparability of the groups.

Beta weights and a multiple correlation were computed with the performance ratings as the dependent variable. The resulting multiple R is an indication of the strength of the relationship between the battery and on-the-job performance. Utilizing the regression equation developed on the validation
group, estimated criterion scores for the subjects in the cross-validation group were computed. The procedure was followed to show the validity of the battery when the effects of working with optimum weights are eliminated.

Tenure as a criterion measure was studied apart from the validation project based on additional information available in the subjects' personnel folders. Police officers initially employed during the stipulated period who left the Bureau were separated into two categories: those who voluntarily left the Bureau although in good standing and those who were considered to be performing unsatisfactorily. T-tests of differences in mean scores and F-tests for differences in standard deviations were computed where differences appeared to approach significance for those who left versus those who stayed and between the two categories of officers who left.

Statement of Hypotheses

The total study can be considered a test of four basic hypotheses:

Hypothesis I

Written test scores, interview scores and psychological scores should be significantly correlated with job performance.

Hypothesis II

The total selection battery should be more predictive
of job performance than any one of the instruments taken individually.

Hypothesis III

Police officers who elected to stay with the Bureau should have significantly higher mean scores on each predictor variable than those who leave.

Hypothesis IV

Ex-police officers whose performance was satisfactory should have significantly higher mean scores on the predictors than ex-officers whose performance was considered to be unsatisfactory.
CHAPTER IV

RESULTS

An analysis of scatter diagrams showing the correlation of each predictor variable with the criterion variable revealed that the trend of the relationship was linear and homogeneous throughout the range. Therefore, linearity and homoscedasticity, the two basic assumptions for zero order product moment correlations, were met.

A comparison of the variable score distributions in the validation sample and the cross-validation sample substantiated that random assignment had resulted in two almost identical groups (see Table II).

In the matrix of intercorrelations for the validation group, the psychological variable was found to have the strongest relationship with job performance (see Table III). The written test variable, positive as in the case of the psychological variable, was second strongest in magnitude. The interview variable had no appreciable relationship, either positive or negative, with the criterion variable.

A similar matrix was computed for the cross-validation sample in order to observe stability of the correlations between the samples. In the cross-validation sample, the coefficient for the psychological variable was found to be
<table>
<thead>
<tr>
<th>Validation Sample</th>
<th>Cross-Validation Sample</th>
<th>F-ratio/T-Value</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Written Test</td>
<td>Written Test</td>
<td>t = -0.21</td>
<td>N.S.</td>
</tr>
<tr>
<td>( \bar{X} = 96.57 )</td>
<td>( \bar{X} = 96.78 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD = 6.56</td>
<td>SD = 6.77</td>
<td>F = 1.06</td>
<td>N.S.</td>
</tr>
<tr>
<td>Interview</td>
<td>Interview</td>
<td>t = 0.73</td>
<td>N.S.</td>
</tr>
<tr>
<td>( \bar{X} = 53.63 )</td>
<td>( \bar{X} = 52.90 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD = 7.08</td>
<td>SD = 6.68</td>
<td>F = 1.12</td>
<td>N.S.</td>
</tr>
<tr>
<td>Psychological</td>
<td>Psychological</td>
<td>t = 0.57</td>
<td>N.S.</td>
</tr>
<tr>
<td>( \bar{X} = 83.66 )</td>
<td>( \bar{X} = 83.34 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD = 3.78</td>
<td>SD = 3.74</td>
<td>F = 1.02</td>
<td>N.S.</td>
</tr>
<tr>
<td>Performance</td>
<td>Performance</td>
<td>t = 0.11</td>
<td>N.S.</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Evaluation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \bar{X} = 50.00 )</td>
<td>( \bar{X} = 49.85 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD = 9.42</td>
<td>SD = 9.13</td>
<td>F = 1.06</td>
<td>N.S.</td>
</tr>
</tbody>
</table>

N.S. = Not Significant
TABLE III

INTERCORRELATION MATRICES

**Validation Group (N = 137)**

<table>
<thead>
<tr>
<th></th>
<th>Written Test</th>
<th>Interview</th>
<th>Psychological</th>
<th>Performance Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Written Test</td>
<td>——</td>
<td>0.091</td>
<td>0.198**</td>
<td>0.125</td>
</tr>
<tr>
<td>Interview</td>
<td>——</td>
<td>——</td>
<td>0.266***</td>
<td>0.002</td>
</tr>
<tr>
<td>Psychological</td>
<td>——</td>
<td>——</td>
<td>——</td>
<td>0.145*</td>
</tr>
<tr>
<td>Performance</td>
<td>——</td>
<td>——</td>
<td>——</td>
<td>——</td>
</tr>
</tbody>
</table>

**Cross-Validation Group (N = 68)**

<table>
<thead>
<tr>
<th></th>
<th>Written Test</th>
<th>Interview</th>
<th>Psychological</th>
<th>Performance Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Written Test</td>
<td>——</td>
<td>0.146</td>
<td>0.122</td>
<td>0.213***</td>
</tr>
<tr>
<td>Interview</td>
<td>——</td>
<td>——</td>
<td>0.294**</td>
<td>0.132</td>
</tr>
<tr>
<td>Psychological</td>
<td>——</td>
<td>——</td>
<td>——</td>
<td>0.062</td>
</tr>
<tr>
<td>Performance</td>
<td>——</td>
<td>——</td>
<td>——</td>
<td>——</td>
</tr>
</tbody>
</table>

* Significant at .10 level
** Significant at .05 level
*** Significant at .01 level
lower while the coefficients for the written test and inter-
view were higher. Based on the two matrices there seems to
be present a true relationship between the written test and
the psychological with job performance. It is interesting
to note that in both groups the largest coefficient is
between the interview variable and the psychological variable.

The correlation between the three predictor variables
and the criterion variable by means of multi-regression
resulted in a primary validation index of 0.18 which is
significant at the 0.05 level of confidence (Table IV).
The multiple R for the hold-out group when computed for
comparison purposes was found to be 0.24.

To provide additional information concerning the
variable interrelationships the written test, interview and
psychological were each in turn treated as a dependent
variable (Table V). When related to the other three vari-
ables the psychological shows a consistent and relatively
strong relationship with the interview and to a lesser de-
gree with the written test.

A cross-validation coefficient of 0.12 was obtained
when the weights derived by the multi-regression analysis
on the validation group were applied to predictor scores in
the hold-out group. Because the weight for the interview
variable in the primary validation group, being negative,
was a potential source of instability the predicted per-
formance scores for the subjects in the cross-validation
### TABLE IV

**BETA VALUES AND MULTIPLE R's**

#### Validation Group (N = 137)

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Written Test</th>
<th>Interview</th>
<th>Psychological</th>
<th>Performance</th>
<th>Mult. R.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Written Test</td>
<td>—</td>
<td>0.04</td>
<td>0.17</td>
<td>0.10</td>
<td>0.22</td>
</tr>
<tr>
<td>Interview</td>
<td>0.04</td>
<td>—</td>
<td>0.25</td>
<td>-0.04</td>
<td>0.27</td>
</tr>
<tr>
<td>Psychological</td>
<td>0.16</td>
<td>0.25</td>
<td>—</td>
<td>0.12</td>
<td>0.34</td>
</tr>
<tr>
<td>Performance</td>
<td>0.10</td>
<td>-0.04</td>
<td>0.14</td>
<td>—</td>
<td>0.18</td>
</tr>
</tbody>
</table>

#### Cross-Validation Group (N = 68)

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Written Test</th>
<th>Interview</th>
<th>Psychological</th>
<th>Performance</th>
<th>Mult. R.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Written Test</td>
<td>—</td>
<td>0.10</td>
<td>0.08</td>
<td>0.20</td>
<td>0.26</td>
</tr>
<tr>
<td>Interview</td>
<td>0.09</td>
<td>—</td>
<td>0.28</td>
<td>0.09</td>
<td>0.33</td>
</tr>
<tr>
<td>Psychological</td>
<td>0.08</td>
<td>0.28</td>
<td>—</td>
<td>0.01</td>
<td>0.30</td>
</tr>
<tr>
<td>Performance</td>
<td>0.20</td>
<td>0.10</td>
<td>0.01</td>
<td>—</td>
<td>0.24</td>
</tr>
</tbody>
</table>
TABLE V

PREDICTOR VARIABLE MEANS AND STANDARD DEVIATIONS FOR FOUR GROUPS OF SUBJECTS BASED ON EMPLOYMENT STATUS

<table>
<thead>
<tr>
<th></th>
<th>Written Scores</th>
<th>Interview</th>
<th>Psychological</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current and have been promoted one level (N = 38)</td>
<td>X = 101.0*</td>
<td>X = 54.5</td>
<td>X = 85.2</td>
</tr>
<tr>
<td></td>
<td>SD = 6.8</td>
<td>SD = 8.3</td>
<td>SD = 4.2</td>
</tr>
<tr>
<td>Non-Failure terminations (N = 27)</td>
<td>X = 97.9</td>
<td>X = 55.1</td>
<td>X = 84.9</td>
</tr>
<tr>
<td></td>
<td>SD = 6.6</td>
<td>SD = 7.2</td>
<td>SD = 3.8</td>
</tr>
<tr>
<td>Failure terminations (N = 34)</td>
<td>X = 97.6</td>
<td>X = 53.8</td>
<td>X = 84.8</td>
</tr>
<tr>
<td></td>
<td>SD = 11.8*</td>
<td>SD = 6.3</td>
<td>SD = 5.2*</td>
</tr>
<tr>
<td>Current and remain at entrance level (N = 150)</td>
<td>X = 97.2</td>
<td>X = 53.8</td>
<td>X = 84.2</td>
</tr>
<tr>
<td></td>
<td>SD = 6.2</td>
<td>SD = 6.6</td>
<td>SD = 3.5</td>
</tr>
</tbody>
</table>

* Differs significantly from current employees who remain at the entrance level
were again computed with the interview variable omitted from the equation. The deletion of the interview variable resulted in an increase in the relationship between the estimated criteria scores and the actual criteria scores to a cross-validation index of 0.16. Neither coefficient, with or without the interview variable, attained statistical significance.

The relationship between the predicted and true criterion is best shown in the form of a scatter diagram (Figure 3). Relative to the higher scores the lower scores seem to be more closely clustered indicating that in the lower range the coefficient could be higher than 0.16.

The hypothesis that the entire battery will be more predictive than any of its component parts is substantiated. However, when beta weights are applied to a new sample of subjects, predictability is diminished; and, due to the sample size of the cross-validation group, does not reach statistical significance.

The differences in variable means and standard deviations for four categories of employees is presented in Table V. Collapsing the scores for police officers who had been promoted into the same pool with the scores for the officers who were classified as currents at the entrance level the means for the written test, interview, and psychological are 97.9, 53.9 and 84.4 respectively. In comparison, the means for the two groups of officers who had terminated
FIGURE 3

SCATTER DIAGRAM SHOWING RELATIONSHIP BETWEEN PREDICTED PERFORMANCE SCORES AND TRUE PERFORMANCE SCORES UTILIZING WRITTEN AND PSYCHOLOGICAL SCORES ONLY
combined are 97.7, 54.4 and 84.8. The differences between the two groupings are not significant. Therefore, the hypothesis that the terminated officers would score significantly lower on the three predictor variables must be rejected.

The written test, interview and psychological mean scores for officers who were non-failure terminations is 97.9, 55.1 and 84.9 respectively. The means for the officers classified as failure terminations is 97.6, 53.8 and 84.9 respectively. As in the case of officers who left versus those who have remained, the differences are insignificant. The difference in interview scores between the two groups is in the hypothesized direction but does not reach statistical significance. Therefore, the hypothesis that the two types of terminations can be distinguished by the predictor variables must be rejected.

The average written test score for officers who have been promoted is significantly greater than the average score for officers who have remained at the entrance level. While this difference is interesting it is not surprising. Promotions to the rank of sergeant or detective are primarily obtained by success on written tests. Except that promotional instruments are more akin to achievement tests than aptitude tests, they are quite similar to the entry test. A high score on the entry test apparently is an indication that a police officer will tend to earn high scores on promotional examinations.
The written test and psychological scores standard deviations for failure terminations are significantly larger than in the other three categories of subjects indicating that those officers whose departure elicited no sorrow tend to differ from the majority of their peers in ways measured by the written test and psychological.
CHAPTER V

DISCUSSION

The relatively low magnitude of the cross-validation coefficient should suggest caution in extolling the virtues of the examination battery but it does not mean that the battery should be hastily abandoned.

The criterion variable was undoubtedly weakened by various common rater tendencies; specifically such tendencies as leniency bias, halo bias, severity bias and error of central tendency. The net affect of these rater biases is to make obscure the degree of relationship between predictor variables and criterion variables. Such unreliability in the criterion variable leads to under-representation of true validity.

Another source of criterion related validity obfuscation is the variation in the amount of knowledge that various supervisor raters have of their respective employees. Some sergeants stay in close working contact with their employees, "covering in" on developing emergency situations while others prefer to work in semi-isolation, only checking with their subordinate employees at roll-call and at the end of each shift. These varying amounts of rater-ratee exposure along with differing preferences as to desirable attributes make
performance ratings a shaky yardstick of true performance at best.

The "criterion problem" is particularly acute in law enforcement validation research because the concept of ideal work performance is elusive. One person's image of the ideal police officer is the aggressive ambitious crime fighter, while another person may prefer as ideal model, the pleasant, non-aggressive father figure who helps small children at street intersections. At times the virtues of utilizing many narrowly defined criterion measures have been extolled, while at other times the global consolidated index of overall proficiency has carried the day. While the work of a police officer is factorially complex, personnel decisions require global evaluations. In spite of the concomitant shortcoming of unreliability an overall index of success or value to the organization is the most meaningful test of selection program validity. However, in evaluating the bivariate intercorrelations, the multiple R, and the cross-validation coefficient, one should keep in mind that performance ratings may reflect factors conceptually unrelated to those intended to be measured by the selection instruments. One officer may rate subjects high who tend to turn in reports without prompting, while another may rate up subjects who always report to work a little earlier than required. The validity coefficient for the written test would undoubtedly have been higher if the criterion instrument had
permitted an evaluation of communication skills, problem solving, and learning ability. A selection program, no matter how comprehensive, cannot cover every attribute that may influence a rating supervisor. While inhouse performance evaluation systems are often characterized by extreme restriction in range due to leniency and other rater tendencies and usually do not relate directly to factors measured by selection processes, they do possess relevance which is often lacking in evaluation systems designed specifically for use in research projects.

Another reason for caution in abandoning the selection program is the restriction in range problem. Since the selection instruments studied were actually utilized in the selection of the subjects, the range of individual differences is necessarily restricted through the attrition process. The markedly severe restriction in range, present in this study, can be seen in the score distributions depicted in Appendices G, H and I. The selection ratio, the proportion of applicants selected to those tested, was 1 to 10.2, 1 to 3.8, and 1 to 1.7 for the written test, interview and psychological respectively. The change in standard deviation for each variable from the total sample tested to the sample actually hired is statistically significant beyond the .01 level of confidence.

The effects of working with restricted ranges can be estimated by formulas designed specifically for that purpose.
The dispersion of scores in the restricted sample and the unrestricted sample in terms of the variable on which selection occurred must be known. The correlation in the restricted sample between the variable to be validated and the criterion variable must also be known. Applying the formula to the coefficient for the written test in the validation group, the corrected coefficient is 0.27 in lieu of 0.12. The estimated correlation for the psychological, assuming that everyone who was interviewed was psychologically assessed, is 0.18 in lieu of 0.14 for the restricted sample in the validation group.

The relationship between the interview and the psychological is of particular interest since it resulted in the largest and apparently the most stable coefficient. Applying the correction formula for restriction in range as if all candidates who were interviewed were given the psychological assessment, the correlation in the passing range of the psychological is 0.39.

It is not unreasonable to speculate that if all candidates had been hired, the resulting multiple R in the validation group and the shrunken R in the cross-validation group would have been in the 0.30 to 0.40 range and the 0.20 to 0.30 range respectively.

The fact that the interview and the psychological show a strong relationship is not too surprising when it is considered that the strongest factor influencing a candidate's
assessment score is probably the psychologist's general impression of the candidate acquired during the clinical interview. The members of the interview panel, although attempting to defer matters of psychological concern, cannot help but be influenced by impressions of a candidate's maturity and emotional stability. This is not to say that one process is the duplication of the other. There is still much variance unaccounted for in both processes. There is some evidence in the table of variable intercorrelations that the psychological may be washing out of the process those candidates who, if appointed, would have permitted the interview to show a stronger relationship with the criterion. Caution is again suggested in arguing that the interview can be eliminated from the process. It is the observation of many that obvious misfits who have survived the objective portion of the entrance examination have, on numerous occasions, been detected and screened out by the oral interview. While the passing interview scores may not relate to the recorded performance levels in a systematic way the interview process may still be effectively screening at the lower end of the score distribution.

The difference in mean scores for four groups of employees shown in Appendices J, K and L are surprising in that it was thought that the predictor variables would show a significant relationship with employee status as hypothesized but, this was obviously not the case. The
significantly higher mean score on the written test for those subjects who later were to obtain promotion to higher levels in the organization was not unanticipated. According to a personal communication from Dr. Terry Eisenberg, then research scientist for the International Association of Chiefs of Police, this phenomenon has been observed in a number of departments. As mentioned earlier, tests tend to be excellent predictors of performance on related tests.

The relatively large written and psychological standard deviations were not anticipated. On intelligence factors measured by the written test and non-intelligence factors measured by the psychological, the officers grouped as failure terminations differ from officers who stay with the department or leave the department in good standing. That the psychological factors involved were non-intelligence in nature was supported by relating the psychological scores to the interview and the written scores. The mean interview score for the 17 failure termination officers with the highest psychological scores was 56.3 while the mean score for the 17 officers with the lowest psychological scores was 51.4. In other words, the failure termination officers who had received high psychological scores, had, as a group, received exceptionally high interview scores. Conversely, terminated officers with low psychological scores had received exceptionally low interview scores. The difference in written scores for the two sub-groups of
failure terminations based on psychological scores was not significant. The conclusion seems to follow that those terminated officers whose performance was unsatisfactory tended to be exceptionally intelligent or exceptionally unintelligent relative to the majority of the police officers hired during the period. Subjects in the failure termination group seemed either to perform exceptionally well or poorly on factors measured by the psychological. These findings are somewhat consistent with Dr. Levy's research findings. While she did not investigate intelligence as such she did find that officers who terminated for cause tended to be exceptionally well educated (1967).
CHAPTER VI

SUMMARY

This study attempted to show the validity of a total police officer selection program and the contribution each major part made to the validity of the whole. The criterion variable consisted of a summary score for each subject reflecting six bi-annual performance evaluations. Two hundred five police officers who had three years of post-probationary employment were randomly assigned to a validation group and a cross-validation group. Beta weights and beta coefficients were developed by multi-regression analysis for each predictor variable utilizing the subjects in the validation group (N-137). The raw scores on each variable for the subjects in the cross-validation group were then fed into the regression equation and predicted performance scores were computed. The predicted scores were then correlated with the obtained scores resulting in a non-significant shrunken $R$ of 0.12. Since it was noted that the interview made a marginal and slightly negative contribution to the multiple predicted performance scores, they were again computed with the interview data excluded. An increase in the shrunken multiple $R$ was anticipated. The multiple $R$ did increase to 0.16 but still did not attain significance.
Reasons for believing that the low cross-validation multiple R's failure to reach significance was due to shortcomings in the criterion variable and the severe restriction in range in the predictor variables were offered. While the use of the multiple hurdle selection format makes corrections for restriction in range difficult, some tentative estimates of the unrestricted coefficients for the written test and psychological were presented. These two predictor variables were selected since they seemed to make the greatest contribution to the multiple R. The estimate for the written test was 0.27 and, for the psychological, 0.14. If the shrunken R of 0.16 is assumed to be the measure of the true relationship between the battery and the criterion with the interview excluded, and is not just a chance artifact, the estimate of shrunken multiple R corrected for restriction of range should fall between 0.20 and 0.30.

A second phase of the study was an analysis of predictor variables within four groups of subjects representing four different categories of employment status; current and promoted, current and not promoted, non-failure terminations, and failure terminations. A hypothesis that the two current groups combined would receive significantly higher scores on the three variables in comparison to the two terminated groups combined was not supported. A hypothesis that failure terminations would have significantly lower scores on the three variables than the non-failure
terminations was not supported. Exceptionally high written test scores for officers who have been promoted and atypically large standard deviations on the written test and psychological for terminated officers who were occupational failures were pointed out and some theoretical explanations for these phenomena were offered.

The characteristics of candidates who after hiring are found to be unsatisfactory in the work role constitutes a promising area of research. In addition to intelligence factors, personality factors seem to loom large as a potential explanation for the inability to adapt to this unique line of work. Hopefully this interesting area of inquiry will begin to receive the attention it deserves.

Hopefully other municipalities will begin to conduct research on total selection battery validation rather than continue with the present piecemeal approach of studying one instrument at a time. Since instruments are never used alone but, are always just one unit in a total selection program, results will always be spurious until instruments are studied in the context of the total system.
REFERENCES


Dunnette, Marvin D. Personnel selection and placement, Belmont, California: Wadsworth, 1966


Narrol, Harvey G. and Levitt, Eugene E. Formal assessment procedures in police selection. Psychological Reports, 1963, 12, 691-694


APPENDIX A

The City of Portland offers non careers with a future through Open Continuous examination for:

POLICE PATROLMAN

Monthly Rate: $633 (starting rate) $653 (after one year) $674 (after two years) $695 (after three years)

TO TAKE THE EXAMINATION YOU MUST:

1. Have passed your 21st birthday by the date of application. Names will be removed from the eligible list when non-veterans reach their 30th birthdays and when veterans reach their 35th birthdays. You will be required to present your birth certificate.

2. Be at least 5'9" tall without shoes and have weight within required proportions to height. Mental or physical defects which would prevent or interfere with effective performance on the job will disqualify. (Each applicant will be responsible for determining that he meets the physical requirements listed on Medical Standard No. 1. Height, weight, and color perception will be checked during the examination process. Successful candidates will be given a complete medical examination before they are appointed.)

3. Have graduated from a standard high school course or have successfully completed the General Educational Development Test with an average standard score of 51 and no single score below 40. Proof of your education will be required before the written test.

4. Have a good moral character and the highest personal integrity. Felony convictions and excessive misdemeanors, including juvenile and traffic offenses, are disqualifying. All arrests for any law violation (including juvenile offenses) must be declared on application. All candidates are fingerprinted and checked through FBI records. Your background and character will be carefully investigated by the Police Bureau before you are appointed.

5. Possess or be able to obtain an Oregon State Driver's License.

THE EXAMINATION WILL CONSIST OF

1. A written test, weighted 60%, designed to measure each candidate's reasoning ability, general intelligence, ability to understand written material and aptitude for police work. No special preparation is necessary or recommended for this test.

2. A qualifying physical agility test, consisting of a series of measures of physical fitness, agility, and endurance.

3. An oral interview, weighted 40%, conducted by an impartial board to assess each candidate's personal qualifications, background and interest in police work.

Before consideration for appointment, candidates successfully completing the above tests will undergo additional intensive screening to further determine their aptitude and suitability for police work.

Failure in any part of the examination process will eliminate the candidate.

FILING APPLICATIONS:

Applications must be filed at the Portland Civil Service Board Room 400, Hughes Building 115 S. W. Fourth Avenue

TIME AND PLACE TO APPEAR FOR WRITTEN EXAMINATION:

Appointment to take the examination will be made at the time of application. Arrangements may be made to take the written test, physical agility test and oral interview on the same day. Evening appointments for the written test are available.

Civil Service Board, Portland, Oregon

Police Patrolman

George R. McDonald, Secretary

Robert L. Ransom, Examiner

Posted: 7-3-67

Annct: 33-67
APPENDIX B

ADDITIONAL INFORMATION FOR APPLICANTS

WHO MAY APPLY: Applicants must possess the minimum qualifications listed in this bulletin and must be citizens of the United States.

EQUAL EMPLOYMENT OPPORTUNITY: All qualified applicants will receive consideration for appointment without regard to race, religion, color, national origin, sex, political affiliations, or any other non-merit factor.

HOW TO APPLY: Applicants must complete and submit an official application form before the closing date. Applications may be submitted by mail if they are postmarked no later than midnight of the closing date. Required documents should be presented with the application and will be returned to the applicant.

HEALTH AND CHARACTER INVESTIGATION: All candidates selected for appointment are required to pass the medical examination prescribed by the Board, and are fingerprinted. All arrests must be declared on the application.

VETERANS PREFERENCE: Upon passing an open competitive examination, an honorably discharged member of the United States armed forces who served during any war for a period of more than 90 consecutive days, or who has served since July 1, 1958 and has received or qualified for the Armed Forces Expeditionary Medal, is granted a preference of 5 points. An additional preference of 5 points is granted to such veteran if he has a service-connected disability of 10% or more. Proof of veteran status or of having received the Expeditionary Medal must be presented within ten days of the date of the written examination. Information on preference for promotional examinations is obtainable at the Civil Service Office.

SALARIES: All salaries as stated are based on present information and are subject to change. Appointments are made at the minimum salary and increases are granted at intervals based on satisfactory service.
# Police Officer Oral Interview

## Appearance, Manner and Bearing

Consider: Will the candidate's physical appearance be a help or hinderance in dealing with the general public? With adult or juvenile offenders? Is the candidate at ease and friendly and yet confident, alert and attentive. Is any hostility or resentment shown toward the interviewers? Is over-confidence demonstrated? Does the candidate appear to have the tact and adaptability necessary to deal with irate citizens, boisterous students, etc., under trying conditions? Would the candidate tend to be too abrupt, submissive, impatient, overbearing?

<table>
<thead>
<tr>
<th>Superior</th>
<th>Satisfactory</th>
<th>Deficient</th>
</tr>
</thead>
</table>

Comments:

## Comprehension and Presentation of Ideas

Consider: Does the candidate grasp ideas quickly? Are responses to questions prompt, clear, orderly and to-the-point? Does the candidate ramble, become confused, vague or indirect? Do you consider the candidate's ability to express thoughts adequate for police work?

<table>
<thead>
<tr>
<th>Superior</th>
<th>Satisfactory</th>
<th>Deficient</th>
</tr>
</thead>
</table>

Comments:

## Attitude Toward Position

Consider: Does the candidate have a real interest in police work or is this just another job? Is there clear evidence to support this interest? Is there an awareness of the hardships and disadvantages of police work? Is the dramatic and glamorous aspects of the work the main attraction? Is the candidate interested in the police profession as a lifetime career?

<table>
<thead>
<tr>
<th>Superior</th>
<th>Satisfactory</th>
<th>Deficient</th>
</tr>
</thead>
</table>

Comments:

## Over-all Capability

Consider: In making your over-all evaluation, consider all of the above factors and any other personal qualifications possessed by the candidate. Do the qualifications fit the job? How good would the Portland Police Bureau be if each Police Officer were of the same caliber as the candidate? To what extent does the candidate meet what you consider to be the ideal qualifications for this job?

<table>
<thead>
<tr>
<th>Superior</th>
<th>Satisfactory</th>
<th>Deficient</th>
</tr>
</thead>
</table>

Comments:
<table>
<thead>
<tr>
<th>POSITION</th>
<th>PERM.</th>
<th>PROB.</th>
<th>TEM.</th>
<th>FROM</th>
<th>TO</th>
<th>PERIOD</th>
<th>BADGE BRN NO.</th>
<th>PRACT. - DIVISION</th>
<th>DATE</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Rate Each Factor</th>
<th>Superior</th>
<th>Competent</th>
<th>Needed</th>
<th>Improvement Needed</th>
<th>Unsatisfactory</th>
</tr>
</thead>
</table>

Use COMMENTS space to describe employer's strengths and weaknesses. Give examples of work well done and plans for improving performance. (Factor ratings of Unsatisfactory or Superior must be substantiated by comments.)

**1. Quantity**
- [ ] Amount of work performed
- [ ] Completion of work on schedule

**2. Quality**
- [ ] Accuracy
- [ ] Thoroughness
- [ ] Effectiveness and neatness of work product
- [ ] Written expression
- [ ] Oral expression

**3. Work Habits**
- [ ] Attendance
- [ ] Compliance with work instructions
- [ ] Compliance with rules, regulations and procedures
- [ ] Application to duties
- [ ] Acceptance of responsibilities

**4. Adaptability**
- [ ] Interest
- [ ] Initiative
- [ ] Performance in new situations
- [ ] Performance in emergencies

**5. Personal Relations**
- [ ] Getting along with fellow employees
- [ ] Meeting and handling the public
- [ ] Personal appearance
- [ ] Respect for authority

**6. Other**

**7. Supervisory Ability** (only for supervisors)
- [ ] Planning and assigning
- [ ] Training and instructing
- [ ] Disciplinary control
- [ ] Evaluating performance
- [ ] Leadership
- [ ] Making decisions
- [ ] Fairness and impartiality
- [ ] Approachability

**OVER-ALL EVALUATION**

<table>
<thead>
<tr>
<th>Unsatisfactory</th>
<th>Improvement Needed</th>
<th>Competent</th>
<th>Superior</th>
</tr>
</thead>
</table>

This report has been discussed with me.

**Employee's Signature**

**Signatures of Reporting Officers**

This report is based on my observation and/or knowledge. It represents my best judgment of the employee's performance.

[ ] I Agree

[ ] I do not recommend the probationer's final and (only for complete appointment, recommendations) complete appointment.

**Rater**

**Reviewer**

**Div. Head**

(Instructions on reverse side)
APPENDIX E

READ THE PERFORMANCE EVALUATION INSTRUCTION MANUAL BEFORE PREPARING THE EVALUATION REPORT. NOTE THE COMPLETED PERFORMANCE REPORTS IN THE BACK OF THE MANUAL.

* * * * * * *

PURPOSE OF THE EVALUATION

Performance evaluation is of prime importance to both the supervisor and the employee and should serve the following purposes:

- Find out if Employee's work is up to standard
- Help employee improve work performance
- Let employee know how he is getting along
- Give recognition for good work
- Determine training needs

The form is divided into seven areas which are called "FACTORS". They are:

1. Quantity
2. Quality
3. Work Habits
4. Adaptability
5. Personal Relations
6. "Other" which has been included to permit the addition of factor(s) important to the job but not included elsewhere in the report
7. Supervisory ability (for Spvs. Only)

* * * * * * *

HOW TO FILL OUT THE EVALUATION REPORT

THE RATING FACTORS

Mark an "X" in one of the four boxes to the right of the "Factor" indicating superior, competent, improvement needed or unsatisfactory.

* * * * * * *

THE USE OF ITEMS

The use of the items listed under factors is optional with the Precinct or Division. In the column of boxes under the factor, indicate strong (+); weak (-); or standard (0), performance on items which you find related to the kind of work done by the employee.

* * * * * * *

THE OVER-ALL EVALUATION

To obtain the "over-all evaluation score", add the numerical factor ratings (using the indicated numerical values), and divide by seven (the number of factors). Enter an "X" in the corresponding "over-all evaluation" box. Flexibility is allowed inasmuch as one category may merit more weight than another.

Forward original report through channels to the Personnel Division. Retain one copy for Precinct or Division 201 file.
DISTINGUISHING FEATURES OF WORK

This is beginning level of police work dealing with patrol, investigation, reporting and arrest.

The incumbent is responsible for assignments in the protection of life and property and in enforcement of laws and ordinances in accordance with departmental instructions and regulations. Employee receives general and special instructions through officers of higher rank who review work methods through personal inspection and discussion. Although assignments are usually in the general field of patrol, employee may be given special assignments.

EXAMPLES OF WORK (Note: These examples are intended only as illustrations of the various types of work performed in positions allocated to this class.)

Patrols a specified beat or district on foot, prowl car, or with the harbor patrol; checks windows and doors in business district, investigates suspicious conditions, gives information both general and specific to inquiries; makes arrests, serves writs, warrants, and subpoenas; checks for parking and other traffic violations, writes tickets, and directs traffic.

Works in plainclothes in vice, juvenile, or intelligence unit; investigates conditions, secures evidence and makes arrests; reports conditions to officers and to Chief.

When assigned to identification unit, takes and classifies fingerprints; searches for fingerprint records; covers scene of crime for latent fingerprints, takes photographs and develops negatives and photostats.

As jailer, performs duties dealing with the security and safety of persons detained in the municipal jail; makes periodic checks of cells and doors; assists in feeding inmates; registers and checks belongings, etc.

NECESSARY KNOWLEDGES, SKILLS, AND ABILITIES (At time of Appointment)

Good general and social intelligence.
Ability to understand and execute oral and written instructions.
Ability to be courteous but firm with the public.
Ability to read and understand laws, ordinances, departmental policies, rules and instructions.
Ability to develop skill in the use of firearms.
Ability to react quickly and calmly under emergency conditions.
Ability to write reports.
Ability to operate an automobile safely and according to traffic laws and rules.
Good powers of observation.
APPENDIX G

DISTRIBUTION OF WRITTEN TEST SCORES AT EACH MAJOR STEP
IN THE SEQUENTIAL SELECTION PROGRAM

1. Written scores of all accepted applicants

2. Written scores of applicants interviewed

3. Written scores of applicants who passed interview

4. Written scores of applicants who were psychologically assessed

5. Written scores of applicants hired
APPENDIX H

DISTRIBUTION OF INTERVIEW SCORES AT EACH MAJOR STEP IN THE SEQUENTIAL SELECTION PROGRAM

1. Interview scores of applicants interviewed

   \[ N = 968 \]
   \[ \bar{X} = 45.9 \]
   \[ SD = 10.0 \]

2. Interview scores of applicants who passed the interview

   \[ N = 549 \]
   \[ \bar{X} = 51.8 \]
   \[ SD = 6.2 \]

3. Interview scores of applicants who were psychologically assessed

   \[ N = 426 \]
   \[ \bar{X} = 52.3 \]
   \[ SD = 6.3 \]

4. Interview scores of applicants hired

   \[ N = 254 \]
   \[ \bar{X} = 53.4 \]
   \[ SD = 6.8 \]
DISTRIBUTION OF PSYCHOLOGICAL SCORES AT EACH MAJOR STEP IN THE SEQUENTIAL SELECTION PROGRAM

1. Psychological scores of all applicants assessed

\[ N = 426 \]
\[ \bar{X} = 82.2 \]
\[ SD = 5.0 \]

2. Psychological scores of applicants hired

\[ N = 254 \]
\[ \bar{X} = 83.8 \]
\[ SD = 4.0 \]
APPENDIX J

DISTRIBUTION OF WRITTEN TEST SCORES FOR FOUR GROUPS OF SUBJECTS BASED ON EMPLOYMENT STATUS

Current and have been promoted at least one level

Non-failure terminations

Failure terminations

Current and remain at entrance level

\[
\begin{align*}
\text{N} & = 38 \\
\bar{X} & = 101.0 \\
\text{SD} & = 6.8 \\
\text{N} & = 24 \\
\bar{X} & = 97.2 \\
\text{SD} & = 6.6 \\
\text{N} & = 34 \\
\bar{X} & = 97.6 \\
\text{SD} & = 11.8 \\
\text{N} & = 150 \\
\bar{X} & = 97.2 \\
\text{SD} & = 6.2
\end{align*}
\]
APPENDIX K

DISTRIBUTION OF INTERVIEW SCORES FOR FOUR GROUPS OF SUBJECTS BASED ON EMPLOYMENT STATUS

Current and have been promoted one level

Non-failure terminations

Failure terminations

Current and remain at entrance level

N = 38
\( \bar{X} = 54.5 \)
SD = 8.3

N = 24
\( \bar{X} = 55.1 \)
SD = 7.2

N = 34
\( \bar{X} = 53.8 \)
SD = 6.3

N = 150
\( \bar{X} = 53.8 \)
SD = 6.6
APPENDIX L

DISTRIBUTION OF PSYCHOLOGICAL SCORES FOR FOUR GROUPS OF SUBJECTS BASED ON EMPLOYMENT STATUS

Current and have been promoted at least one level

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
<th>N</th>
<th>X</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>10</td>
<td>38</td>
<td>85.2</td>
<td>4.2</td>
</tr>
<tr>
<td>85</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Non-failure terminations

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
<th>N</th>
<th>X</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>10</td>
<td>24</td>
<td>84.9</td>
<td>3.8</td>
</tr>
<tr>
<td>85</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Failure terminations

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
<th>N</th>
<th>X</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>10</td>
<td>34</td>
<td>84.8</td>
<td>5.2</td>
</tr>
<tr>
<td>85</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Current and remain at entrance level

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
<th>N</th>
<th>X</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>10</td>
<td>150</td>
<td>84.2</td>
<td>3.5</td>
</tr>
<tr>
<td>85</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## APPENDIX M

### ANALYSIS OF POLICE OFFICER TEST
#### BY QUESTION TYPE

<table>
<thead>
<tr>
<th>Question Type</th>
<th>Number Possible</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Memory section</td>
<td>20</td>
</tr>
<tr>
<td>2. Word definition</td>
<td>34</td>
</tr>
<tr>
<td>3. Math story problems</td>
<td>20</td>
</tr>
<tr>
<td>4. Reading comprehension</td>
<td>22</td>
</tr>
<tr>
<td>5. Hypothetical police situations</td>
<td>8</td>
</tr>
<tr>
<td>6. Analogy</td>
<td>6</td>
</tr>
<tr>
<td>7. General knowledge</td>
<td>10</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>120</strong></td>
</tr>
</tbody>
</table>
APPENDIX N

SUBJECTS' EDUCATIONAL LEVEL ATTAINED AT THE TIME OF APPOINTMENT AS POLICE OFFICERS

<table>
<thead>
<tr>
<th>Level</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>G.E.D. only</td>
<td>19</td>
<td>7</td>
</tr>
<tr>
<td>High school graduation only</td>
<td>97</td>
<td>38</td>
</tr>
<tr>
<td>Up to one year of college</td>
<td>73</td>
<td>29</td>
</tr>
<tr>
<td>Up to two years of college</td>
<td>41</td>
<td>16</td>
</tr>
<tr>
<td>Up to three years of college</td>
<td>15</td>
<td>6</td>
</tr>
<tr>
<td>Up to four years of college</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>254</td>
<td>100</td>
</tr>
</tbody>
</table>
APPENDIX O

AGES OF SUBJECTS AT TIME OF APPOINTMENT
N = 254, \( \bar{X} = 23.9 \) YEARS, S.D. = 2.8 YEARS