1973

Use of Regression Analysis in Police Patrolman Selection, The

J. T. Flynn

M. Peterson

THE USE OF REGRESSION ANALYSIS IN POLICE PATROLMAN SELECTION

J. T. FLYNN AND M. PETERSON

J. T. Flynn Ph.D. is an Associate Professor of Educational Measurement and Evaluation at the University of Connecticut, Storrs, Conn. He is the author of four books in his field of specialization as well as numerous articles which have appeared in various professional journals.

Mayfield Peterson Ph.D. has been Assistant Professor of Educational Psychology at the University of Connecticut. Dr. Peterson has a background in clinical psychology, as well as training in the field of educational psychology and has published several articles in professional journals. In September 1972 he joined the faculty of the Herbert M. Lehman College, the City University of New York.

The history of law enforcement as it relates to metropolitan police departments reveals a continuous change in the concept of what the role and function of metropolitan policemen should be. It is not difficult to understand this continuous change in role and function when we study the industrial and technological changes which have continuously pervaded our society since its inception. Because of industrial and technical developments and the obvious increase in the complexity of our society, it has become necessary in recent years for metropolitan officials to begin to look more carefully at the criteria which have been used over the years for selecting personnel to serve on city police forces. Consequently, for the first time in history many U.S. cities have launched large scale research and experimental projects designed to help municipal officials determine what personal attributes are most suitable to the unceasing demands imposed upon the city policeman in our contemporary technological society.

A perusal of the literature, which is not very replete, suggests that formalized procedures for the selection and training of police recruits are quite new. Only a few decades ago a police officer was considered appropriately selected and adequately trained if he possessed the necessary skills to fire a weapon and to interpret the penal code of his local community. Changes in social structure and operation of cities have made it necessary for city officials to begin to reevaluate procedures for selection and training of city police recruits. This radical change was probably most clearly demonstrated during the years preceding the economic depression of the late nineteen twenties and the early nineteen thirties. It was during this period that defiance of the law and widespread crime became rampant across the nation. During this same period it became quite obvious to city officials that law enforcement agencies were inadequately equipped to cope with the overwhelming amount of crime in our cities. It was during this same period that man was evolving from a horse and buggy era into a motor vehicle form of transportation.

Quite clearly, city and state police departments were left with little or no alternative except to begin to look for new directions for their own operations. Among new directions came selection and training procedures which were quite different from preceding years.

Frost (2) recommended the following as areas worthy of considering in selecting police recruits.

1. A written examination which should consist of three types of testing:
   a. an intelligence test
   b. a police aptitude test
   c. a personality test
2. Physical examination to consist of two parts:
   a. medical-laboratory examination
   b. physical examination
3. Personal interview to be preceded by a complete and accurate investigation of the applicant's character.
4. Personal experience—the final phase.

Of the four stages listed by Frost (2), he recommends that they be followed as listed above and that each succeeding one should be contingent upon success in the preceding stage. That is, applicants not successful in phase one will not be allowed to proceed through stage two.

Adams (1) recommends a battery of written tests to include a test of intelligence, memory, reading comprehension, and spelling. He also recommends

an oral examination, a physical-medical examination, psychiatric examination, polygraph test, and a background investigation.

Osterburg, Trubitt, and Myren (8) in a review of the literature found that selection criteria for Cadet Programs and police recruits are frequently the same. In selecting police cadets and police recruits these researchers found that cities often use tests of mental abilities as well as some measure of mental health. Among tests frequently used were *The California Mental Maturity Test, Minnesota Multiphasic Personality Inventory, Army Alpha, Army General Classification Test,* and the *Otis-Self Administering Test of Mental Ability.*

Mills, McDevitt, and Tonkin (6) found that in 55 U.S. cities having populations greater than 150,000 some sort of psychological testing was used in the selection of police recruits. In addition, 16% of these same cities used some type of psychiatric interview, 85% used some measure of aptitude.

Narrol and Levitt (7) surveyed 61 U.S. cities with populations greater than 150,000 according to the 1950 census. They asked the chiefs of these to supply detailed descriptions of the methods they used to screen candidates for employment as policemen.

Of 61 cities surveyed the authors were able to use results from 55 of the cities. Data from the other cities were inadequate and were eliminated from the studies. These researchers concluded the following from their survey: (a) even those departments which recognize policeman examinations to be intelligence tests, usually use the former in preference to standardized I.Q. tests; (b) little is being done to evaluate the personalities of applicants; (c) it appears that there is little or no research being done upon police selection techniques; and (d) not many psychologists are involved either in research upon, or in the use of, the selection procedures now employed.

McAllister (5) compared performance evaluation of New York City recruits during the first eighteen months of their careers with the initial judgment of the background information gathered on them at the time of their selection. Performance evaluations of the recruits were based on the following criteria: time lost, time lost for injuries, completion of training standards, formal recognition of outstanding performance, absence of formal disciplinary charges, completion of the probationary period, and evaluations by supervisors. The researcher concluded that background investiga-tions were not particularly valid inasmuch as those judged to be of better character did not achieve higher performance than those with fair or poor character evaluations.

Background information as a selection criteria used in isolation from other methods of evaluation can obviously not be considered a very scientific and systematic way of selecting men to serve on police forces. However, we do believe that background information can, if treated thoroughly and systematically, contribute much to the prediction of both quantity and quality of police performance. One major fault in the use of background information that we have observed is that the techniques used are not systematic and orderly and the men who collect the information are not highly skilled in the techniques of information gathering. For example, we have observed that some cases are investigated via personal contacts and extensive interviews while others may be investigated by mail or telephone conversation. In the case of the New York City study we raise the question of how systematic and uniform the system of information gathering was. Were the men investigated by personal contact and interview, by mail, or by telephone, or could it have been a combination of all three? Levy (4), for example did find that biographical data significantly differentiated between successful and non-successful policemen. In an additional study Levy (5) developed a predictive model of tenure for new police officers. She used information from the personnel files of 5,000 men to develop the predictor model. At the time of this study 100 men had terminated their employment in police work for various reasons. Of the 100 men the model predicted sixty-four terminations. The majority of these 100 men (80) were characterized by their employers as failures. The reader should be cognizant of the fact that Levy's predictor study was criticized by Newman who later re-examined the data.

Spencer and Nichols (9) investigated several factors believed to relate to a prediction of police recruit performance. Among these was background information. The authors found a definite relationship between background information and subsequent performance on the job. The study was conducted in the city of Chicago.

It seems clear from the research data based on background investigation that predictions can be made with a considerable degree of confidence from the information gathered. What seems to be cru-

Table 1

Selection Factors	Weights Arbitrarily Assigned
Training and Experiences	1
PPA Test Score	4
Oral Evaluation	5

cial, as we see it, is that the investigation must be thorough, systematic, objective and that the same information must be gathered on each applicant using the same techniques of data collection for each candidate.

Purpose: This study was designed to evaluate the existing procedures for police recruit selections in a metropolitan northeastern city having a population of nearly 200,000. From a measurement point of view, it seems futile to speculate about what types of selection criteria are most helpful in determining who is most likely to become successful police officers. Rigorous social scientific strategies are available that allow systematic investigations into the feasibility and efficiency of any series of selection procedures that might be currently utilized. Consequently, a study was designed and conducted to determine the empirical relationships that might exist between conventional selection procedures and subsequent performance in a Police Academy training program. The project was designed to analyze the screening procedures used for police applicants, the test used in selection, and its contribution to the criterion of performance during training, along with two other selection indices, background and training and an oral evaluation interview.

Description of the Selection Procedure: In the city's selection procedure, each candidate received a score, Training and Experience, based on his background, training, and experiences related to police work. In addition to this score, each candidate was administered a test, the Public Personnel Association (PPA) test. This is an instrument designed to assess the extent to which a person possesses the skills necessary for successful performance as police officers. Further, each applicant was required to be interviewed by an oral evaluation board, comprised of members of the Police and Personnel Departments. A numerical score representing the consensus of this board was used as the third criterion score for selection purposes.

The three indices were then combined and an overall index, a General Average, was computed

for each man. If this score exceeded the cut-off score, the man was selected. It was at this point that an analysis of the selection procedures revealed that an arbitrary weighting process was being used regarding the three selection factors, oral evaluation, the PPA test, and the Training and Experience score. These weights are shown in Table 1.

This strategy was employed to reflect the rationale that the behaviors assessed by the Oral Evaluation are of primary importance in terms of their contribution to the criterion of policeman performance. In other words, it was weighted five times as heavily as the Training and Experience factor.

It is important to note that this type of procedure for assigning weights to varying types of information gathered for the purpose of selection is commonly used. The weights used are most frequently generated from a logical and rational point of view. However, since the Means and Variances are expressed in unequal units of measurement, they cannot be considered in any comparative sense. Simply multiplying each candidate's score on the factors by the respective weight will not reflect the appropriate weights in the total, or in this case, General Average score.

Method: The records of thirty nine recruits, covering a period of six months, including the recruit class currently in progress, were utilized. Scores for the three selection factors, Training and Experience, PPA Test, and Oral Evaluation, were recorded for each subject. Further, each recruit's final grade, received upon completion of training, was recorded. This measurement was used as the criterion score for the analysis. It should be indicated that more suitable criteria were discussed, such as some supervisory ratings of patrolmen's performance. However, they were not available to be incorporated into the study. This condition, then, resulted in the study being based upon an important assumption, which was open to question. It was assumed, that a positive, linear relationship existed between performance in training for recruits and their subsequent performance as patrolmen. Within reasonable limits this assumption probably holds. The extent to which it did not hold represented a clear limitation of the study.

Results The PPA test for the 39 subjects in this study produced a mean score of 85.6 (70 = minimum passing score) with a standard deviation of 10.7.

One significant index of the usefulness of any

test is its reliability. This concept suggests how stable, over a period of time a test is. The group of 39 recruits were administered the PPA Test a second time, approximately four months after a first administration. The reliability index, $r = .32$, is a statistically significant one, but a consideration of the amount of variance attributable from one test administration to the other, approximately 10%, provides evidence that the test is a highly unreliable instrument, and offers no consistent set of measurements. A review of the test manual indicated an absence of technical information regarding standardization, norms, reliability and validity.

Since the PPA Test is a predictor instrument, designed to aid in accurately selecting the best risks for potential police work, it was decided to analyze the performance of a group of experienced police officers on the test. A group of 23 officers, with a minimum of two years on the force were administered the test. Item difficulty and discrimination indices were reviewed for the test performance of this group. Table 2 presents descriptive data on the test.

The t statistic was utilized to test the significance of the difference between the Mean Scores for the Recruits and Police Officers. The difference, 7.53 points, yielded a t value that was very highly significant. This indicates that, as a group, the recruits performed significantly higher than the police officer group. This information takes on meaning when it is recognized that the characteristics measured by the PPA test are those thought to be crucial for effective functioning as a policeman. In view of this, it was reasonable to predict that experienced police officers would perform better than recruits. In this instance, they did not.

Some additional information emerged regarding the comparison of discrimination and difficulty indices for both the Police Officer and the Recruit groups. Table 3 presents these data.

Apparently, the PPA Test provides little in the way of discriminative power for both recruits and experienced police officers.

In addition to the review of the PPA Test, and its use as a selection instrument, attention was directed to an analysis of the relationship between the other selection procedures, Training and Experience and the Oral examination indices. The ensuing discussion of these procedures, when viewed as a selection strategy, is based upon the following information regarding the correlations

Table 2

Means and Standard Deviation of Police Officer Scores on the Police Patrolman Test PPA

No.	$\bar{X}$	S.D.	Scores	
			Hi	Lo
23	78.09	17.40	102	35

between each predictor and the criterion training performance.

$$r_{1.2} = .191 \qquad r_{2.3} = .116$$
$$r_{1.3} = .588 \qquad r_{2.4} = .301$$
$$r_{1.4} = .528 \qquad r_{3.4} = .345$$

Where 1 = T&E Score
      2 = PPA Test
      3 = Oral Eval.
      4 = Final Grades

It is important to direct attention to the fact that the Training and Experience variable, by itself, correlates most significantly, $r = .528$, with the criterion variable. In turn, the PPA correlates, $r = .301$, and the Oral Examination, $r = .345$. This clearly indicates that both the test and the oral examination procedure do not add any significant information to the total selection arrangement. Stated another way, the best single predictor of performance in training is the individual's score on the Experience and Training variable.

The above discussion treats each of the three selection procedures independently of each other. Current evidence in the behavioral sciences indicates that identifiable variables do not operate independently, but rather, interact upon each other in unique ways to produce the final behavioral outcome.

To identify the parameters of the data gathered

Table 3

Discrimination and Difficulty Indices for the PPA Test for Police Officers and Recruits[1]

	Discrimination		Difficulty	
	Good	Poor	Good	Poor
Officers	56%	44%	43%	57%
Recruits	51	37	51	43

[1] For the purposes of the investigation, the range of values .40 to .60 was established for good, difficulty, and discrimination indices.

Table 4

For 3 Variables Entered
Multiple Correlation Coefficient, R = 0.568
Adjusted for degrees of freedom R = 0.534
Regression Coefficients, Variable 1, .781
Variable 2, .143
Variable 3, .032

more clearly, they were subjected to a statistical procedure, based upon the Step-Wise, Multiple Regression Model. This technique allows us to gain insight into how the relevant predicators, taken singly and together in all possible combinations, contribute to the final outcome, prediction of performance scores on the criterion.

The basic logic of multiple regression is illustrated in a regression formula, predicted variable $(Y) = a + bX_1 + bX_2 \ldots + bX$ where, in this instance, Y is the performance in training, and X_1 is Training, X_2 is the PPA Test score, and X_3 is the Oral Examination score. These represent actual weights that provide insight into relative contributions to the criterion. This technique was used in order to take into account the compensatory nature of the data available. By this is meant that it should be possible for persons who score low in one area to make up that low performance by scoring high in another area. Multiple regression allows this to happen. Techniques such as multiple cut-off scores and discriminant function analysis are non-compensatory in nature.

A computer analysis of the data, utilizing a step-wise, multiple regression model revealed the information presented in Table 4.

Table 4 provides cumulative data encompassing what has been discussed in this report to clearly indicate that while the Training and Experience Score (.78) contributes the most to the criterion performance, both the PPA Test (.14) and the Oral Examination Score (.03) contribute somewhat less to the criterion. Ordinarily, since the PPA Test and the Oral score are uncorrelated, $r = .116$, they both should provide increased efficiency in selecting candidates. Since the increase in the Multiple Coefficient, from .52 to .56 is so slight when these two predictors are used, it suggests that they are not tapping what is required to perform well in training. The data yielded the following prediction formula, including the constant value and the respective data weights:

$$Y = A + bX_1 + bX_2 + bX_3$$

or

$$Y = 16.53 + .78X_1 + 14X_2 + .03X_x$$

This regression analysis takes into account the relative contribution of each of the selection factors, and its use allows a specific grade to be predicted for each candidate who has measurements on the three selection factors.

Discussion: This investigation lends empirical support to the notion that logical, a priori choices of selection criteria for police candidates can be very misleading. In this instance, for example, the use of the Oral Evaluation and the PPA Test was not warranted in light of data showing that they provided little if any additional information beyond the Training and Experience information.

The Step-Wise Multiple Regression Analysis is an extremely powerful and useful technique appropriate for analyses of selection factors utilized by Police Departments. As many as five selection factors can be included in any attempt to predict successful police candidates. The procedure can reveal factors that are not providing useful information in final selection, and can point out additional factors that might be more beneficial in efficiently selecting the best personnel.

It is suggested that there is a need to better scrutinize available predictors in order to generate relevant variables that are uncorrelated with each other, and positively correlated with the predicted criterion behavior. If this is accomplished, the accuracy and efficiency of selection procedures can be significantly improved by isolating meaningful predictors and avoiding duplication. It might be noted that the use of correlated predictors for selection is a clear redundancy, since correlated measures, to a large extent, measure the same behavior.

This experiment serves to demonstrate the usefulness of Step-Wise Multiple Regression for personnel selection. The technique offers the advantage of changing the analyses of data that is typically collected for the purposes of selection. The technique does not require that new information be collected, but simply that information available for applicants be analyzed in a different and more powerful manner. Additionally, the technique offers the advantages of (1) being objective and resistant to bias, (2) compensatory and fair to all applicants, (3) realistic regarding the nature of human behavior as we know it, (4) defensible as a selection device established in a systematic man-

ner, and (5) potentially capable of aiding in the ultimate recognition of what types of behavior directly contribute to successful personnel job performance.

A final observation which may be of value to officials concerned with police recruit selection is the finding that experience and training represented the best single predictor for subsequent performance in the academy. Since experience and training represent, at least in part, background information, then this component might become crucial as a predictor of future success in the academy. It is particularly interesting that this finding is similar to the data published by Levy (4) and Spencer and Nichols (9). It is therefore recommended that very refined and sophisticated models be developed and used for purposes of gathering background information. Any system of gathering background information on police recruit applicants must be used uniformly and consistently. Every applicant must be investigated via the same system of information gathering. It will not suffice to investigate some cases by intensive personal interview, others by mail, and still others by telephone conversations. It has been our experience that inconsistency in information gathering has

been the rule rather than the exception in the screening of police recruit applicants which usually results in unreliable information.

REFERENCES

1. ADAMS, THOMAS F., LAW ENFORCEMENT: AN INTRODUCTION TO THE POLICE ROLE IN THE COMMUNITY, 22–36 (1968).
2. FROST, THOMAS M., A FORWARD LOOK IN POLICE EDUCATION, 4–5 (1959).
3. Levy, Ruth J., *Predicting Police Failures* 58 J CRIM L, CRIMINOLOGY AND POL SCI 265–276 (1967).
4. Levy, Ruth J., *Investigation of a Method for Identification of the High Risk Police Applicant*, Institute for Local Self Government, Berkeley, Calif. (1970).
5. McAllister, John A., *A Study of Prediction of Measurement of Police Performance*, 14 POLICE, 58–64 (1970).
6. Mills, Robert, R. J. McDevitt, and Sandra Tonkin, *Situational Tests in Metropolitan Police Recruit Selection*, 57 J. CRIM L, CRIMINOLOGY AND POL SCI, 99–106 (1966).
7. Narrol, Harvey G. and Eugene E. Levitt, *Formal Assessment Procedures in Police Selection*, 12 PSYCHOLOGICAL REPORTS, 691–694 (1963).
8. Osterburg, James W., Hieland, J. Trubitt, and Richard A. Myren, *Cadet Programs: An Innovative Change*, 58 J CRIM L, CRIMINOLOGY AND POL SCI, 112–118 (1967).
9. Spencer, Gilmore and Robert Nichols, *A Study of Chicago Police Recruits*, 38 POLICE CHIEF, 50–55 (1971).