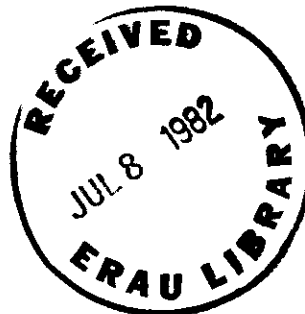


UNCLASSIFIED

Doc  
FAA  
AM  
79  
14

**DTIC**



# Technical Report

distributed by



**Defense Technical Information Center  
DEFENSE LOGISTICS AGENCY**

Cameron Station • Alexandria, Virginia 22314

UNCLASSIFIED

# NOTICE

We are pleased to supply this document in response to your request.

The acquisition of technical reports, notes, memorandums, etc., is an active, ongoing program at the Defense Technical Information Center (DTIC) that depends, in part, on the efforts and interests of users and contributors.

Therefore, if you know of the existence of any significant reports, etc., that are not in the DTIC collection, we would appreciate receiving copies or information related to their sources and availability.

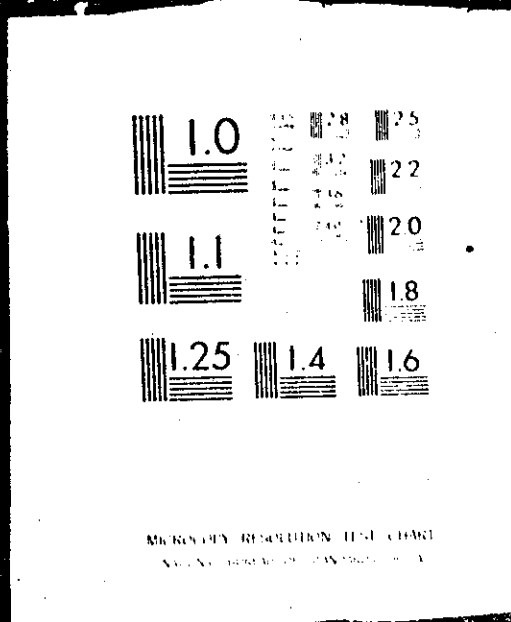
The appropriate regulations are Department of Defense Instruction 5100.38, Defense Technical Information Center for Scientific and Technical Information (DTIC); Department of Defense Instruction 5129.43, Assignment of Functions for the Defense Scientific and Technical Information Program; Department of Defense Directive 5200.20, Distribution Statements on Technical Documents; Military Standard (MIL-STD) 847-A, Format Requirements for Scientific and Technical Reports Prepared by or for the Department of Defense; Department of Defense Regulation 5200.1-R, Information Security Program Regulation.

Our Acquisition Section, DTIC-DDA-1, will assist in resolving any questions you may have. Telephone numbers of that office are: (202) 274-6847, 274-6874 or Autovon 284-6847, 284-6874.

1 OF 1

AD

A068581



AD A068581

DDC FILE COPY

14

FAA-AM-79-14

# LEVEL

41

6

THE SELECTION OF AIR TRAFFIC CONTROL SPECIALISTS:  
TWO STUDIES DEMONSTRATING METHODS TO INSURE AN  
ACCURATE VALIDITY COEFFICIENT FOR SELECTION DEVICES

10

James O. Boone  
Mary A. Lewis

10

12p.

Civil Aeromedical Institute  
Federal Aviation Administration  
Oklahoma City, Oklahoma



DDC  
RECEIVED  
MAY 15 1979

11  
MAR 1979

Document is available to the public through the  
National Technical Information Service  
Springfield, Virginia 22161

Prepared for  
U.S. DEPARTMENT OF TRANSPORTATION  
Federal Aviation Administration  
Office of Aviation Medicine  
Washington, D.C. 20591

264320

LB

NOTICE

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its content or use thereof.

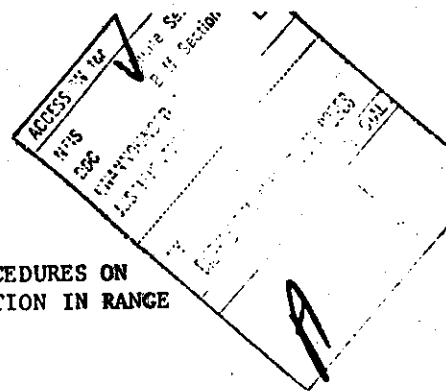
Technical Report Documentation Page

1. Report No. FAA-AM-79-14		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle THE SELECTION OF AIR TRAFFIC CONTROL SPECIALISTS: TWO STUDIES DEMONSTRATING METHODS TO INSURE AN ACCURATE VALIDITY COEFFICIENT FOR SELECTION DEVICES				5. Report Date MARCH 1979	
				6. Performing Organization Code	
7. Author(s) James O. Boone and Mary A. Lewis				8. Performing Organization Report No.	
9. Performing Organization Name and Address FAA Civil Aeromedical Institute P.O. Box 25082 Oklahoma City, Oklahoma 73125				10. Work Unit No. (TRAIS)	
				11. Contract or Grant No.	
12. Sponsoring Agency Name and Address Office of Aviation Medicine Federal Aviation Administration 800 Independence Avenue, S.W. Washington, D.C. 20591				13. Type of Report and Period Covered	
				14. Sponsoring Agency Code	
15. Supplementary Notes Work was performed under Tasks AM-C-78-PSY-66 and 70.					
16. Abstract There are several conditions that can influence the calculation of the statistical validity of a test battery such as that used to select Air Traffic Control Specialists. Two conditions of prime importance to statistical validity are recruitment procedures and the accuracy of the data base. The recent edition (1978) of the Federal Uniform Guidelines on Employee Selection Procedures places considerable emphasis on recruitment practices and their effect on validity. In the first of two studies, Monte Carlo techniques were employed to demonstrate the frequently overlooked effect that recruitment procedures can have on the validity coefficient. It was shown how highly specific recruitment results in a more homogeneous group of applicants which produces a small applicant group variance on the selection test scores. It was further shown how a small applicant group variance considerably reduces the validity coefficient when the coefficient is corrected for selection effects; commonly termed restriction in range. This paper suggests a procedure that eliminates this recruitment problem and results in compliance with the Uniform Guidelines. The second study describes a statistical procedure to use when it is necessary to eliminate erroneous data. The procedure employs the notions of statistical distance and probability to identify data that has an extremely small likelihood of belonging to the population of the remaining data set.					
17. Key Words Selection Recruitment Restriction in Range Outliers Multivariate Distance Functions				18. Distribution Statement Document is available to the public through the National Technical Information Service, Springfield, Virginia 22161	
19. Security Classif. (of this report) Unclassified		20. Security Classif. (of this page) Unclassified		21. No. of Pages 10	22. Price

415

AN EXAMPLE OF THE EFFECTS OF RECRUITMENT PROCEDURES ON  
CORRECTING THE VALIDITY COEFFICIENT FOR RESTRICTION IN RANGE

James O. Boone and Mary A. Lewis



I. Introduction.

Fundamental to the selection of Air Traffic Control Specialists (ATCS) are the recruitment procedures used to attract job applicants. Although frequently overlooked, different approaches to recruitment can readily affect the statistical assessment--i.e., the statistical validity measure--of the tests or devices used to qualify or rank ATCS applicants for job consideration.

Recently renewed interest in the validity coefficient can be attributed, to some degree, to the adoption of the Uniform Guidelines on Employee Selection Procedures (7) by the Equal Employment Opportunity Commission (EEOC), the U.S. Civil Service Commission (CSC), the Department of Labor, and the Department of Justice. Since the four agencies adopting the guidelines are charged ultimately with insuring equitable practices in selection and other employment decisions, for both private industry and Federal and state agencies, their adoption of the guidelines has the effect of establishing them as a standard for all government and private organizations. The guidelines elaborate on the technical standards and the size of validity coefficients for validation of selection devices. As a result of the guidelines, the validity coefficient is of prime interest to employers in terms of selection, placement, and promotion.

It has long been recognized that the size of a correlation coefficient is affected by the range or variance of the measures being correlated (2,4,6). The selection test scores of persons who have already been selected for a given type of position are a more homogeneous set of measures than the scores from the applicant group. When this more homogeneous set of measures is correlated with a criterion of job success, a smaller validity coefficient is obtained than would be produced by using the original and larger applicant group's selection test scores. In a study related to selecting pilot trainees, Thorndike (6) demonstrated that selection (in his case 13 percent of the applicants were selected) can produce a rather drastic reduction in the validity coefficient; one of the coefficients actually changed from a .40 to -.03. Given the Uniform Guidelines' emphasis on the validity coefficient, it is understandable why employers are interested in correcting the validity coefficient for this restriction in range due to selection.

Thorndike (6), Gulliksen (2), and others have given various formulas to correct the validity coefficient for restriction in range. However, the appropriate use of these correction formulas has been the source of some

discussion. While there is a general agreement in the literature that extreme selection poses a considerable threat to the accuracy of the corrections (1, 3,5), there have been questions about violating the assumptions underlying the formulas (1).

The purpose of this paper is to investigate a frequently ignored issue in selection research that can have a sizeable effect on the correction of a validity coefficient for restriction in range. The present study will explore an example of the effects of recruitment styles on the magnitude of the validity coefficient that has been corrected for restriction in range and suggest one method to help minimize these undesirable effects.

## II. An Example.

Suppose, for example, that two companies, or agencies, A and B, each hired 50 persons over a period of time to perform essentially the same job. The same selection test was employed by both companies. As a standard practice, company A maintained a general ad in the local newspaper and, when persons responded, they related to the respondents what jobs were available and then tested those applicants who were interested. Company B, however, had a different recruitment policy. Company B advertised specific jobs, stating specific qualifications that must be met prior to the applicant's being tested. In both companies the applicant groups and the hired groups were proportional to the available work force population in terms of race and sex. Both companies performed a validity study and corrected the validity coefficients for restriction in range.

In the situation described above, company A will have tested a group of applicants with a wider range of abilities and consequently will have a considerably larger variance among their applicants' test scores than will company B. The research question to be answered by the present study is: What effect do these recruitment styles, and their resulting applicant group variances, have on the corrected validity coefficient? In order to answer this question, several different unrestricted, or applicant group variances were used in the correction formula with the restricted, or selected group variance held constant to determine the effect of the different unrestricted variances on the magnitude of the validity coefficient.

## III. Methods.

The formula used to correct for restriction in range in the present study is Thorndike's formula 6 (ref. 6, p. 173) or its equivalent, Gulliksen's formula 18 (ref. 2, p. 137):

$$RR_{xy} = \frac{R_{xy} \cdot \frac{SS_x}{S_x}}{\sqrt{1 - R_{xy}^2 + R_{xy}^2 \frac{SS_x^2}{S_x^2}}} \quad (1)$$



where  $SSx^2$  = the applicant group's test variance,  $Sx^2$  = the selected group's test variance,  $Rxy$  = the correlation between the selected group's test scores and a criterion of job success, and  $RRxy$  = the estimated correlation between the applicant group's test scores and a criterion of job success. The difference between the variance on variable x for the applicant group ( $SSx^2$ ) and the selected group ( $Sx^2$ ) is used in the formula to represent the amount of restriction in variance due to selection on variable x.

To demonstrate the effect of using different applicant group variances to correct the validity coefficient, the following procedure was employed. Formula 1 was used with the ratio  $SSx/Sx$  varied from 3.0 to 2.5 to 2.0 to 1.5.  $RRxy$  was then estimated by formula 1 while varying  $Rxy$  from .01 to 1.00 in increments of .01.

#### IV. Results.

Figure 1 demonstrates the effects of using different unrestricted variances in the correction formula. The  $RRxy$  estimates are plotted as a function of  $Rxy$  for each of the four unrestricted variances.

Table 1 shows the mean  $RRxy$  estimates for each of the four unrestricted variances and the standard deviations of the estimates. The means were computed by converting the correlations to Fisher's z.

TABLE 1. Means for the Estimates of  $RRxy$  for Each of the Four  $SSx/Sx$  Ratios

<u><math>SSx/Sx</math> Ratios</u>	<u>Means of <math>RRxy</math> Estimates</u>
1.5	0.605
2.0	0.672
2.5	0.709
3.0	0.755

#### V. Discussion.

It is clear from Figure 1 and Table 1 that as the  $SSx/Sx$  ratio becomes larger, the magnitude of the corrected validity coefficient also increases. As the values of  $Rxy$  move toward the middle values, the discrepancies between the estimated validity coefficients for the different unrestricted variances become even more pronounced. To extend the hypothetical situation to the given example, if company A had an  $SSx/Sx$  ratio of 3.0, and company B had an  $SSx/Sx$  ratio of 1.5, as illustrated in Figure 1, at an  $Rxy$  value of .10, which is a practical value for an explicitly restricted correlation, the corrected validity coefficient would be .14 for company B and .26 for company A. The increase in applicant variability resulted in an estimated correlation

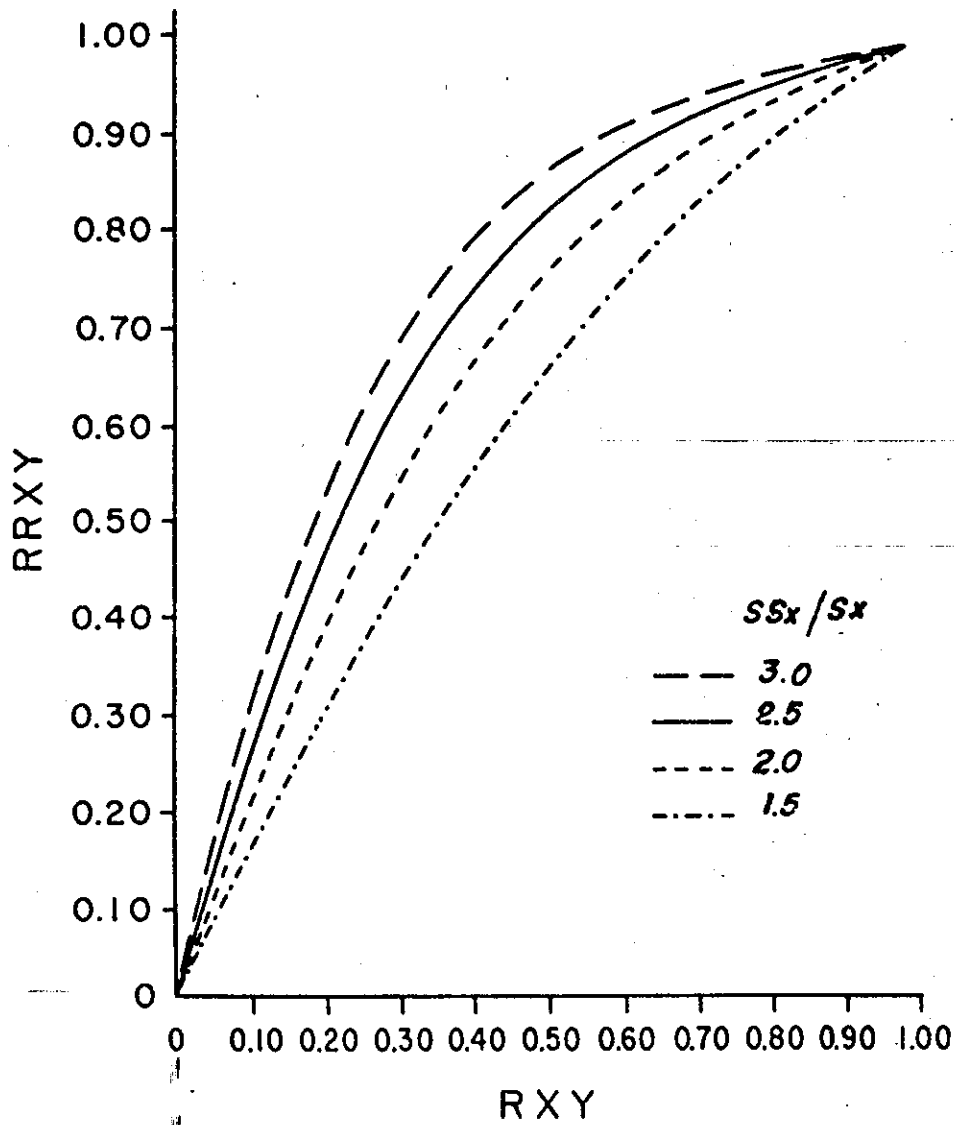


Figure 1. The effects of recruitment on correcting for restriction in range in the two-variable case.

for company A that was almost twice that for company B even though the persons selected for the jobs were the same.

The stipulation in the Uniform Guidelines concerning the sample used for validity studies is that the sample ". . . should insofar as feasible include the racial, ethnic, and sex groups normally available. . .". However, the convention of using the applicant variance as an estimate of unrestricted variance is not a guideline nor a necessity. For companies which prefer to recruit by advertising in a highly specific manner, one solution would be to obtain the unrestricted test variance by administering the selection instrument to other applicants regardless of what job they are seeking. This method would help alleviate the restricting effect of recruitment procedures, since the variance used in the correction formula would not have been restricted as much by recruitment. This procedure would also aid in appropriately maximizing the corrected validity coefficient, because the estimated unrestricted validity coefficient would be a better generalization to the available labor market, which is a requirement of the Uniform Guidelines. Failure to use appropriate correction techniques can result in an underestimate of the validity of selection tests and may leave a company vulnerable to divergent interpretations under the Uniform Guidelines.

#### REFERENCES

1. Brewer, J. K., and J. R. Hills: Univariate Selection: The Effect of Size of Correlation, Degree of Skew, and Degree of Restriction, PSYCHOMETRIKA, 34, 1959.
2. Gulliksen, H.: Theory of Mental Tests, New York, Wiley and Sons, 1950.
3. Lord, F. N., and M. R. Novick: Statistical Theories of Mental Test Scores, Reading, Massachusetts, Addison-Wesley Company, 1968.
4. Pearson, K.: Mathematical Contributions to the Theory of Evolution--XI. On the Influence of Natural Selection on the Variability and Correlation of Organs, Philosophical Transactions, 200-A, pp. 1-66, 1903.
5. Rydberg, S.: Bias in Prediction: On Correction Methods, Stockholm, Almqvist and Wiksells, 1963.
6. Thorndike, R. L.: Personnel Selection, New York, Wiley and Sons, 1949.
7. Uniform Guidelines on Employee Selection Procedures: Federal Register, Vol. 43, No. 166, August 25, 1978.

A STATISTICAL PROCEDURE FOR ELIMINATING EXTREME, DEVIANT SCORES FROM  
THE LONGITUDINAL AIR TRAFFIC CONTROL DATA BASE

James O. Boone

I. Introduction.

With large files of data it is to be expected that some of the columns of data will contain inaccuracies. For example, on a multiple choice test occasionally some individuals mark the same option for every item. Others may mark the same option for several items in a row consistently throughout the test. These types of problems are easily eliminated by inspection of the answer sheets. Another possible source of inaccuracies is data input errors. Each column of data is manually input by hand and carefully crosschecked; however, data input errors may still occur. In the case of input error, if a score lies outside the range of possible scores for that test, it can readily be seen and corrected. Inaccuracies of the type listed above are usually detected by close inspection.

There are other situations where inaccuracies can occur that cannot be detected by inspection. For example, inaccurate data inputs that are within the range of possible test scores cannot be identified by inspection. There is what is termed the "christmas tree" effect, where a person simply goes down the answer sheet and marks options at random. Another example occurs when a person answers the first few items appropriately and then gets out of sequence by one item in marking the remaining items. All of these situations can affect the accuracy of the data, while producing scores that are within the range of possible scores. Inaccurate data of this type cannot be identified by inspection.

To comply with U.S. Civil Service Commission (CSC) requirements in eliminating data that is not an obvious error, an appropriate statistical procedure and criterion must be employed. Removal of erroneous data in the ATCS longitudinal data base by means of an appropriate statistical procedure and criterion is the concern of this paper.

II. Method.

The general idea of eliminating extreme, deviant scores that appear to belong to a different population than the remaining data involves the development of a reasonable criterion or rule for score elimination. The following procedures employ the notion of distance and probability to develop a rule for eliminating extreme, deviant scores.

In the univariate case it is assumed that the scores are a random sample from a normally distributed population. The data is transformed to standard form by:

$$X' = \frac{X - \bar{X}}{S}$$

The  $X'$  score would then be a measure of the score's distance from the distribution mean. However, assuming the score in question is an extreme deviant from the distribution mean, then the score's large deviation would bias the computation of the mean and standard deviation. In order to compensate for this effect, the score being evaluated is removed from the data prior to the computation of the mean and standard deviation and then evaluated in terms of its distance from the distribution of the remaining scores. The  $X'$  in question is evaluated by referring to the well-known normal probability function and the probability that  $X'$  belongs to the distribution of the remaining scores is determined. By a preestablished probability criterion, the score in question is either eliminated or maintained as a part of the data. This procedure is repeated for each score.

In the multivariate case, it is assumed that the scores are a random sample from a multivariate normally distributed population and the univariate case is generalized to multivariate space. The multivariate mean or centroid and variance-covariance matrix is computed without the case that is being evaluated, and then the distance and probability are computed as in the univariate case.

The generalized distance function is given in matrix notation by:

$$D = \{(\underline{X} - \bar{\underline{X}})' S^{-1} (\underline{X} - \bar{\underline{X}})\}^{1/2},$$

where  $\underline{X}$  = a score vector,  $\bar{\underline{X}}$  = the vector of means, and  $S$  = the dispersion matrix. This expression is equivalent to Mahalanobis'  $d$  statistic. (The reader is spared the laborious task of going through the derivations to arrive at the multivariate distance function; however, a concise presentation of the Mahalanobis derivation appears in Cooley and Lohnes (2).) It should be noted that  $\bar{\underline{X}}$  and  $S$  are computed without the score vector of the case being evaluated.

The probability function in the multivariate situation can be shown to be distributed as the well-known  $F$ :

$$F = \frac{n-p-1}{p} \cdot \frac{D}{1-\frac{D}{n}}$$

(Again, the reader is spared the derivations; however, Anderson (1) has a clear description.) If the probability associated with the calculated  $F$  exceeds the preestablished criterion, the case is eliminated. This procedure, as in the univariate situation, is repeated for each vector of scores.

### III. Discussion.

The most important consideration in using this procedure is the establishment of the probability criterion for eliminating scores. The purpose of the procedure is to eliminate inaccurate scores. Elimination of deviant scores that are true low or high scores would serve only to decrease the validity correlation between the selection tests and the criterion of job success.

Consider the formula for a Pearson Product Moment (PPM) correlation:

$$r_{xy} = \frac{1}{n} \frac{\sum xy}{\sigma_x \sigma_y} \quad (4)$$

Eliminating inaccurate deviant scores on the average would decrease the individual measures of variation,  $\sigma_x$  and  $\sigma_y$ , without a proportional decrease in the covariation of X and Y,  $\sum xy$ . However, eliminating true low or high scores that predict well would on the average decrease the covariation of X and Y,  $\sum xy$ , without a proportional decrease in their individual variations,  $\sigma_x$  and  $\sigma_y$ . This would result in a spuriously lowered validity coefficient.

Setting the probability criterion is a judgment. The primary consideration in this judgment should be the sample size. In large samples it is more probable that large deviant scores are accurate values. In small samples there is less opportunity for true large deviant scores, and consequently, large deviant scores are less probable. For example, in a random sample of 1,000 one would want to eliminate scores that have a probability of less than 1 in 1,000 ( $p = .001$ ) of belonging to the population represented by the remaining scores. For a sample size of 50, however, one would not want to eliminate a score with a probability of less than 1 in 50 ( $p = .02$ ). This would be too liberal since the elimination of scores is based on the probability that the score belongs to the population of the remaining scores. A random sample of 1,000 would be representative enough of the population to establish a direct relationship between sample size and the probability criterion. A sample size of 50, though, would not on the average be representative of the population. A  $p = .01$  or  $.005$  would be more appropriate for a sample size of 50.

In the case of the current ATCS data, the sample size is approximately 2,000. Consequently, in using the above-described procedures, a probability of  $p = .0005$  can be reasonably set as the criterion. This criterion and the above procedures in general should meet the CSC requirements as a reasonable statistical procedure for eliminating inaccurate data.

#### REFERENCES

1. Anderson, T. W.: An Introduction to Multivariate Statistical Analysis, New York, Wiley, 1958.
2. Cooley, W. W., and P. R. Lohnes: Multivariate Data Analysis, New York, Wiley, 1971.