

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI[®]

Csp

629. Thesis - 13
670 Range

A TEST OF THE RELATIVE AND INCREMENTAL PREDICTIVE
VALIDITY ASSOCIATED WITH A SET OF PAPER-AND-PENCIL
TEST MEASURES AND A SET OF SITUATIONAL TEST
MEASURES WITHIN AN ASSESSMENT CENTRE

by: Leonard W. Slivinski

Thesis presented to the School of Graduate Studies
of the University of Ottawa as partial fulfilment
of the requirements for the degree of Doctor
of Philosophy



Ottawa, Canada, 1978

Université d'Ottawa
BIBLIOTHÈQUES



LIBRARIES
University of Ottawa

UMI Number: DC52490

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform DC52490
Copyright 2007 by ProQuest LLC
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT OF

A TEST OF THE RELATIVE AND INCREMENTAL PREDICTIVE
VALIDITY ASSOCIATED WITH A SET OF PAPER-AND-PENCIL
TEST MEASURES AND A SET OF SITUATIONAL TEST
MEASURES WITHIN AN ASSESSMENT CENTRE¹

A sample of two hundred and ninety-seven subjects from lower middle to upper middle levels of management within the Canadian Federal Government participated in an assessment centre program designed to facilitate the identification of potential for effective functioning at executive levels of management. During the assessment program, performance was assessed through a series of situational tests and a battery of psychometric paper-and-pencil tests. The general progression achieved by these participants approximately one and one-half to four years following assessment was evaluated by means of criterion measures related to position-level and salary, as well as a composite measure combining both of these criterion variables. The research was primarily designed to compare the predictive validity of a set of paper-and-pencil test measures with the validity associated with a set of situational test measures. As part of this effort, the research also focussed on the incremental validity which could ensue from the combination of both sets of predictors. The results of the investigation indicated that there

¹ Leonard W. Slivinski, doctoral thesis presented to the School of Psychology of the University of Ottawa, Ontario, August, 1978, VIII - 167 p.

was no significant difference in the amount of predictive variance accounted for by the paper-and-pencil versus situational testing methods. It was determined, however, that both sets of predictors added in an incremental fashion in the overall prediction of progression within the organization. The existence of this relationship was to some extent dependent upon the criterion measure considered.

ACKNOWLEDGEMENTS

This thesis was prepared under the supervision of Professor Charles E. McInnis, Ph.D., Associate Professor in the School of Psychology. The writer gratefully acknowledges his advice, direction, and interest in the research.

The co-operation of the staff of the Public Service Commission of Canada was appreciated for enlisting the support of all personnel who participated in and made possible the completion of the research.

CURRICULUM STUDIORUM

Leonard W. Slivinski was born November 15, 1941 in Vancouver, British Columbia. He received his Bachelor of Social Science degree from St. Patrick's College, Ottawa, Ontario, in 1965. He received his Master of Arts degree in Psychology from the University of Ottawa, Ottawa, Ontario in 1971. The title of his thesis was: "Divergent Production Profiles and Occupational Membership".

TABLE OF CONTENTS

Chapter	page
INTRODUCTION	vii
1 - REVIEW OF THE LITERATURE AND STATEMENT OF THE PROBLEM	1
1. Psychometric Paper-and-Pencil Testing	5
2. Assessment Centre Prediction of Managerial Success	17
3. Situational Versus Paper-and-Pencil Components of an Assessment Centre	23
4. Summary and Statement of the Basic Hypotheses	36
2 - EXPERIMENTAL DESIGN	39
1. Subjects	39
2. The Context of the Assessment Centre Program	41
3. The Measuring Instruments	43
4. Procedures-Testing, Scoring, Data Gathering	57
5. Statistical Analysis	60
3 - PRESENTATION AND DISCUSSION OF RESULTS	64
1. Descriptive Statistics and Intercorrelation Matrices for the Paper-and-Pencil Predictors, the Situational Predictors, and the Criterion Measures	64
2. Statistical Testing of the Null Hypothesis	74
SUMMARY AND CONCLUSIONS	86
BIBLIOGRAPHY	97
Appendix	
I. Frequency and Percentage Distributions for Relevant Subject Variables	107
II. Description of the Sub-Tests of the General Intelligence Test	117
III. Description of the Dimensions of the Personality Research Form	122
IV. Conversion Equivalency Table for Various Classification Groups	128

TABLE OF CONTENTS

Appendix	page
V. Frequency and Percentage Distribution for the Predictor and Criterion Measures	131
VI. The Multiple R's and Associated Regression Weights Obtained when the Paper-and-Pencil Test Measures and the Situational Test Measures are Related to the Criterion Measures	158
VII. Abstract of: A Test of the Relative and Incremental Predictive Validity Associated with a Set of Paper- and-Pencil Test Measures and a Set of Situational Test Measures Within an Assessment Centre	165

LIST OF TABLES

Table	page
I.- Means, Standard Deviations and Measures of Kurtosis and Skewness for the Nine Paper-and-Pencil Predictors	65
II.- Intercorrelation Matrix of the Nine Paper-and-Pencil Predictors	66
III.- Means, Standard Deviations, and Measures of Kurtosis and Skewness for the Fourteen Situational Predictors	68
IV.- Intercorrelation Matrix of the Fourteen Situational Predictors	69
V.- Correlation of the Nine Paper-and-Pencil Predictors with the Fourteen Situational Predictors	71
VI.- Means, Standard Deviations, Intercorrelation Matrix, and Measures of Kurtosis and Skewness for the Three Criterion Measures	72
VII.- Pearson r's and Multiple R's Relating the Nine Paper-and-Pencil Predictors to the Three Criterion Measures	75
VIII.- Pearson r's and Multiple R's Relating the Fourteen Situational Predictors to the Three Criterion Measures	76
IX.- Multiple R's and Statistical Tests of Differences Obtained in Relating Paper-and-Pencil Predictors and Situational Predictors to Three Criterion Measures .	78
X.- Multiple R's and F Test Differences Obtained With and Without the Addition of Paper-and-Pencil Predictors	81
XI.- Multiple R's and F Test Differences Obtained With and Without the Addition of Situational Predictors	83

INTRODUCTION

Within the area of industrial-organizational psychology, there has been a rapid growth over the past two decades in the use of the assessment centre approach to identify, select and develop managerial talent. As originally conceived within military and organizational settings, the assessment centre approach represents a multi-method evaluative process consisting of at least two prime assessment procedures. One is the more traditional psychometric battery of paper-and-pencil tests, the other is the newly emerging situational or behavioural testing method. Within the assessment centre context, this latter procedure entails the observation of behaviour in a variety of role-playing exercises designed to simulate job-related managerial situations and functions.

Although some research has been generated on the predictive validity of the assessment centre approach as a whole, very little information is available regarding the relative predictive validity of the paper-and-pencil and situational components of the process. The few investigations which have attempted to compare these two assessment methods have been marred by methodological and interpretative difficulties, leading to equivocal results and the need for further systematic investigation.

The purpose of the present research is to establish the relative predictive validity of situational and paper-and-pencil tests within the context of an assessment centre utilized for the evaluation of senior executive potential in the Federal Government of Canada.

As part of this effort, an attempt was made to determine the incremental predictive validity which ensues through the combination of information from both assessment procedures.

The first chapter of this study presents a review of the pertinent literature. It offers a brief overview of the paper-and-pencil and situational testing methods. This is followed by a global review of the predictive validity of psychometric paper-and-pencil tests in evaluating managerial potential. Similar consideration is given to the predictive validity of the assessment centre approach. Researches concerned with the specific comparison of the paper-and-pencil and the situational components of the assessment centre are then presented, concluding with a statement of the null hypotheses relevant to this research.

The second chapter describes the subjects, test instruments and experimental procedures utilized in this research project, along with the statistical analyses employed to test the null hypotheses.

The final chapter presents and discusses the statistical results, leading to a summary of the findings and suggestions for further study.

CHAPTER 1

REVIEW OF THE LITERATURE AND STATEMENT OF THE PROBLEM

The common focus of most psychometric paper-and-pencil instruments has been to generate quantifiable information regarding the traits, characteristics or structures which underlie human behaviour. Basic to this psychometric approach to assessment is the notion of "psychic determinism",^{1,2} whereby the behaviour of an individual is seen as being motivated or determined by the various underlying traits or hypothetical constructs. Moreover, implicit in this conceptualization of human functioning is an assumption that such underlying factors determine an individual's actions across a wide variety of situations.³ The identification and understanding of these psychological characteristics assumedly allows for the prediction of behaviour "...independent of the situation in which the person is immersed."⁴

1 Marvin R. Goldfried and Ronald N. Kent, "Traditional Versus Behavioural Personality Assessment: A Comparison of Methodological and Theoretical Assumptions", in Psychological Bulletin, Vol. 77, 1972, p. 409 - 420.

2 Leona E. Tyler, "Design for a Hopeful Psychology", in American Psychologist, Vol. 28, 1973, p. 1020 - 1029.

3 Goldfried and Kent, Op. Cit., 1972, p. 410.

4. Donald N. Bersoff, "Silk Purses into Sow's Ears - The Decline of Psychological Testing and a Suggestion for its Redemption", in American Psychologist, Vol. 28, 1973, p. 892.

Particularly over the past decade the assumptions underlying the psychometric approach to assessment have come under questioning primarily because insufficient consideration is given to the dynamic aspect of individuals who are constantly interacting differentially to a multitude of environmental and situational variables. Several critical reviews have been written in this area with some authors^{5,6} expressing criticism of the general psychometric tradition while others have focused more specifically on the psychometric aspects of intelligence testing^{7,8} and personality assessment.^{9,10} As an alternate approach to assessment and prediction, each of these authors encourage the use of behavioural or situational testing, as exemplified by such methods as direct observation of real-life performance and indirect observation through various role-playing simulations. In contrast with the psychometric approach of focusing on the traits a person "has", situational testing emphasises what a person "does" in

5 Bersoff, Op. Cit., 1973, p. 892 - 899.

6 Paul F. Wernimont and John P. Campbell, "Signs, Samples and Criteria", in Journal of Applied Psychology, Vol. 52, 1968, p. 372 - 376.

7 David C. McClelland, "Testing for Competence Rather than Intelligence", in American Psychologist, Vol. 28, 1973, p. 1 - 14.

8 Alexander G. Wesman, "Intelligent Testing", in American Psychologist, Vol. 23, 1968, p. 267 - 274.

9 Goldfried and Kent, Op. Cit., 1972, p. 409 - 420.

10 W. Mischel, Personality and Assessment, New York, Wiley, 1968.

various situations. Rather than hypothesising underlying constructs, emphasis rests with the observation of an individual's responses to specific aspects of the environment.¹¹

In the realm of industrial-organizational psychology, this contrast or conflict between psychometric paper-and-pencil and situational testing methods has come to the fore within the evaluative process known as the assessment centre. Originally developed within military¹² and organizational settings,¹³ the assessment centre is viewed as a multi-method assessment approach incorporating both psychometric paper-and-pencil measures and situational tests. Other assessment techniques frequently used within an assessment centre include selection interviews, projective techniques, and the use of biographic and career-related information. The paper-and-pencil instruments typically include diverse measures of cognitive abilities and aptitudes, interests, values, and personality. The situational component involves the observation of behaviour in a number of individual and group role-playing exercises designed to be representative and reflective of a variety of job-related situations. The

11 Goldfried and Kent, Op. Cit., 1972, p. 412.

12 Office of Strategic Services (OSS) Assessment Staff, Assessment of Men, New York, Rinehart, 1948.

13 Douglas W. Bray and Donald L. Grant, "The Assessment Centre in the Measurement of Potential for Business Management", in Psychological Monographs, Vol. 80, 1966, Whole No. 625, p. 1 - 27.

rationale for the employment of this diversity of assessment procedures is to arrive at a "holistic" picture of the individual within the organizational context, and to provide a frame for making "sufficiently" reliable predictions of a person's usefulness to the organization.¹⁴

To date the major use of the assessment centre process within organizations has been directed towards the identification and prediction of managerial potential. Extensive research has been generated to support the overall validity of this multi-method evaluation approach as an aid in managerial selection. However, very little research evidence is available relevant to the contribution made by two major assessment centre components, namely psychometric paper-and-pencil measures and situational tests.

The intent of this initial chapter is to review the literature as it relates to this issue. The initial section will concentrate on a general evaluation of the role of paper-and-pencil tests in the prediction of managerial performance. This will be followed by a similar global review of the predictive validation of the assessment centre process. A third section will concentrate on the specific research to date in which attempts have been made to determine the contribution made by the psychometric and situational testing components of the assessment centre process. The final section will present an overall summary of the problem and the working hypotheses.

14 OSS, Op. Cit., 1948.

1. Psychometric Paper-and-Pencil Testing

Traditional psychometric prediction of managerial effectiveness within industrial-organizational settings has entailed the assigning of scores to test instruments and relating these in some statistical manner to scores on a criterion measure.¹⁵ In much of the early research related to the identification of managerial talent, emphasis appeared to rest with establishing and isolating the single or individual predictor dimensions which could be related to effectiveness and success within an organization. Beginning in the fifties¹⁶ and extending throughout the sixties, increasing consideration was given to the use of multiple predictors tapping various aspects of human performance (e.g. personality dynamics, interest profiles, personal values and attitudes, biographic and career history, cognitive capacity, and so forth) in order to arrive a multi-dimensional view of the individual and of his potential value within the organizational context. Throughout this period more emphasis was placed on the incremental validity which followed through the actuarial or clinical combination of multiple predictors of managerial performance. A scan of the major management journals since the late sixties suggests that psychometric research of this nature has almost come to a halt, being

15 Abraham K. Korman, "The Prediction of Managerial Performance: A Review", in Personnel Psychology, Vol. 21, 1968, p. 296.

16 R. Taft, "Multiple Methods of Personality Assessment", in Psychological Bulletin, Vol. 56, 1959, p. 333 - 352.

supplanted to a large extent by various writings, testimonials, and investigations of the assessment centre approach. The following two sections make a somewhat arbitrary distinction between researches emphasising the validity of individual predictors and those emphasising multiple prediction of managerial performance.

a. Individual Predictors of Managerial Success:

Research directed towards the identification of various singular or unidimensional predictors of managerial potential has been extensively summarized in various reviews, with the most global and comprehensive effort being made by Ghiselli in a continuing series of articles and texts.^{17,18,19} The most recent summary in 1973 consists of a review of all published, and many unpublished reports pertaining to the occupational validity of tests covering the period from 1920 through 1971. As part of this review, Ghiselli differentiates two levels of managerial occupations, first-line supervisors and executives or administrators, and presents the mean validity coefficients of all tests classified into nineteen categories. In relating these categorized test scores to a criterion of success, usually in the form of ratings, the majority of the mean validity coefficients were in the

17 Edwin E. Ghiselli, The Measurement of Occupational Aptitude, Los Angeles, University of California Publications, 1955.

18 Edwin E. Ghiselli, The Validity of Occupational Aptitude Tests, New York, Wiley, 1966.

19 Edwin E. Ghiselli, "The Validity of Aptitude Tests in Personnel Selection", in Personnel Psychology, Vol. 26, 1973, p. 461 - 477.

range extending from the low to high twenties. Overall Ghiselli concludes:

Tests of intellectual abilities, spatial and mechanical abilities, and perceptual accuracy tend to be the best, and are of moderate validity for both types of managerial occupations. Tests of motor abilities have lesser, though apparently some, validity for proficiency criteria. Measures of personality and interest also are of moderate value in predicting the level of proficiency executives and administrators attain on their jobs, but they are much less useful for foremen.²⁰

Other reviews of various aspects of the research literature have arrived at conclusions which are both similar and dissimilar to those of Ghiselli. For example, Korman's review²¹ of the individual predictive validity of various paper-and-pencil instruments resulted in findings not completely in agreement with Ghiselli. In this effort, Korman reviewed mostly published literature of managerial prediction extending from 1947 through 1966; he concerned himself with those studies which were clearly predictive in nature, not considering investigations which utilized a concurrent validity paradigm. Similar to Ghiselli he found that measures of cognitive ability, particularly verbal ability, were "fair" predictors of performance for first line supervisory personnel; however, unlike Ghiselli's conclusions, such measures were not found to be good predictors of higher level

²⁰ Ibid., p. 467 - 468.

²¹ Korman, Op. Cit., 1968, p. 295 - 322.

managerial performance. As well, measures stemming from various personality and interest inventories were generally not shown to have predictive validity for either managerial level, with the possible exception of one projective test.

Three additional reviews have specifically concentrated on this latter area of paper-and-pencil measures of personality and interest. Guion and Gottier²² summarized the pertinent research from two major journals covering a twelve year period (1952 - 1963), concentrating on both personality and interest tests used for selection to managerial and non-managerial occupations. Throughout the review the authors express varying degrees of pessimism tempered with some optimism. In part, they state:

Of greatest importance, it must be concluded that, taken as a whole, there is no generalizable evidence that personality (and interest) measures can be recommended as good or practical tools for employee selection... The best that can be said is that in some situations, for some purposes, some personality (and interest) measures can offer helpful predictions.²³

In 1965 Nash²⁴ reviewed all of the literature to that date relevant to the validity of standard interest inventories for predicting managerial performance. As he did not differentiate managers according to level, direct comparison with the conclusions of

22 Robert M. Guion and Richard F. Gottier, "Validity of Personality Measures in Personnel Selection", in Personnel Psychology, Vol. 18, 1965, p. 135 - 164.

23 Ibid., p. 159.

24 Allan N. Nash, "Vocational Interests of Effective Managers, A Review of the Literature", in Personnel Psychology, Vol. 18, 1965, p. 21 - 37.

Ghiselli in this area is not possible. Overall, Nash found "...a thread of consistent evidence..."²⁵ to support the idea that there is a pattern of interests which distinguishes the more effective from the less effective manager.

Finally, a review by Hedlund²⁶ concentrated on the research related to one specific personality inventory, the Minnesota Multiphasic Personality Inventory. He found practically no evidence to support its use within industrial-organizational settings. Campbell, Dunnette, Lawler and Weick²⁷ have since commented that perhaps this is not unexpected for a test whose development is based on psychological aberration.

Taken as a whole the evidence emerging from these reviews prohibits the establishment of any firm concluding statements. Although the global reviews of Ghiselli point out the potential role of diverse paper-and-pencil instruments in predicting managerial success, the conclusions of other reviews serve to point out some areas

25 Ibid., p. 34

26 Dalia E. Hedlund, "A Review of the MMPI in Industry", in Psychological Reports, Vol. 17, 1965, p. 875 - 889.

27 John P. Campbell, Marvin D. Dunnette, Edward E. Lawler, and Karl E. Weick, Managerial Behaviour, Performance, and Effectiveness, New York, McGraw Hill, 1970, p. 138.

of disagreement. The one major conclusion which has been drawn by various authors^{28,29,30} is that if predictive validity is established for a certain test or dimension, the resulting coefficient tends not to be of high magnitude. For example, in using the mean validity coefficient as an index measure, the statistical data presented by Ghiselli³¹ indicates that approximately ten percent of the predictor-criterion variance is accounted for by measures of intelligence, or personality, or interest. In fact, it is this state of affairs which led Dunnette³² to propose in the early sixties that a moratorium be established on the development of further psychometric tests until those already available were better utilized.

b. Multiple Prediction of Managerial Success:

One of the difficulties or limitations associated with many of these paper and pencil predictive studies is that the emphasis has rested on the individual predictive validity of a test instrument or of a subtest within an inventory instead of the possible incremental contribution that a test might make to a multiple prediction of managerial effectiveness. Guion and Gottier in their summary of the

28 Ghiselli, Op. Cit., 1973, p. 461 - 477.

29 Guion and Gottier, Op. Cit., 1965, p. 135 - 164.

30 Korman, Op. Cit., 1968, p. 295 - 322.

31 Ghiselli, Op. Cit., 1973, p. 461 - 477.

32 Marvin D. Dunnette, "Personnel Management", Annual Review of Psychology, Vol. 8, 1962, p. 285 - 314.

research acknowledge "...that the contribution of a particular test to multiple prediction has been totally overlooked".³³ Ghiselli as well notes that only the individual contribution of tests was considered in his review. He is impressed with the individual predictive validities obtained, in part because "...judiciously selected combinations of tests..." would probably have resulted in higher validity.³⁴

In this light, the review of Campbell, Dunnette, Lawler and Weick³⁵ in 1970 is of interest for their focus was on multiple psychometric assessment programs avoiding studies carried out on small or limited samples or under conditions which seemed to limit generality of the findings. Among the major studies evaluated were those of Standard Oil of New Jersey, the University of Minnesota, American Chamber of Commerce, Stanford University, Sears-Roebuck, and so forth. A common element in all these investigations was the exploration of several psychometric techniques including personality, interest, and biographic inventories, management judgment tests, and measures of cognitive abilities. Conclusions were based on an analysis of the pattern of correlations, their replicability and generalizability, as well as on the incremental predictive validity which ensues from the combination of the various paper-and-pencil test data.

33 Guion and Gottier, Op. Cit., 1965, p. 139.

34 Ghiselli, Op.Cit., 1973, p. 477.

35 Campbell et al, Op. Cit., 1970, p. 148 - 198.

As one important example mention may be made of the efforts in Standard Oil of New Jersey, reported in a series of articles by Laurent.^{36,37,38} In a research project concerned with the Early Identification of Management Potential,³⁹ Laurent employed a wide variety of paper-and-pencil tests reflective of multiple dimensions of performance. Most of these tests were empirically keyed using a concurrent validation strategy within a total sample of 433 managers representative of various levels of management. Several test instruments were shown to contribute to the overall multiple prediction of managerial success; these included two measures of cognitive ability (Miller Analogies Test and the Non-Verbal Reasoning Ability Test), a personality inventory known as the Guilford Zimmerman Temperament Survey, a paper and pencil measure of managerial judgment, a personal history survey, and a self-performance report.

In a later study of Laurent,⁴⁰ the individual correlations of some of the test instruments are given for the original sample, as well as for three additional European samples. The coefficients of

36 Harry Laurent, "Early Identification of Managers", in Management Record, May 1962, p. 33 - 38.

37 Harry Laurent, "Research on the Identification of Management Potential", in John A. Myers (Ed.), Predicting Managerial Success, Ann Arbor, Braun and Brumfield, 1968.

38 Harry Laurent, "Cross-Cultural Cross-Validation of Empirically Validated Tests", in Journal of Applied Psychology, Vol. 54, 1970, p. 417 - 423.

39 Laurent, Op. Cit., 1962, p. 33 - 38.

40 Laurent, Op. Cit., 1970, p. 417 - 423.

correlation of the biographic survey to an overall composite index of managerial success ranged from .44 to .65. For the test of management judgment the correlations extended from .34 to .52 across the four samples; for the personality inventory, the range was from .27 to .33. In terms of the overall multiple correlation for the test battery as a whole, the coefficients ranged from .70 to .79,⁴¹ accounting for approximately fifty percent of the variance. Other studies reported by Laurent,⁴² using shortened versions of the battery for use with nonmanagement personnel, have generally reported validities in the .50's.

In summarizing these and several other research investigations conducted throughout the fifties and sixties, Campbell et al. conclude:

Taken together, these studies provide good evidence that a fairly sizable portion (30 to 50 percent) of the variance in general managerial effectiveness can be expressed in terms of personal qualities measured by self-response tests and inventories and combined by predetermined rules or statistical equations.⁴³

A scan of the major management journals since the time of this comprehensive review indicates that with the exception of two studies

41 Some of these latter coefficients were cited in Campbell et al., Op. Cit., 1970, p. 170.

42 Laurent, Op. Cit., 1968.

43 Campbell et al., Op. Cit., 1970, p. 195.

by Grimsley and Jarrett,^{44,45} few research investigations have been reported on the validation of multiple psychometric paper-and-pencil tests for managerial effectiveness. Much of the available information has centered on the role of paper-and-pencil instruments within the assessment centre approach, the results of which will be reviewed in a later section.

Grimsley and Jarrett attempted to determine the extent to which measures of mental ability, personality, interest, and values differentiated top from middle managers. Using a concurrent validity strategy with a sample of two hundred managers from various companies, they employed a matched design in order to control for the moderating effects of age, as well as of field and level of education. The paper-and-pencil test inventories included the Employee Aptitude Survey, the Gordon Personal Inventory, the Gordon Personal Profile, the Guilford Zimmerman Temperament Survey, the Kuder Preference Record, and the Study of Values. Various combinations of the test data resulted in multiple correlations extending from .52 to .74. Within these various analyses, components of all the test instruments, particularly those of cognitive ability and personality, were found to contribute in an incremental fashion to the overall multiple correlation.

44 Glen Grimsley and Hilton F. Jarrett, "The Relation of Past Managerial Achievement to Test Measures Obtained in the Employment Situation: Methodology and Results", in Personnel Psychology, Vol. 26, 1973, p. 31 - 48.

45 Glen Grimsley and Hilton F. Jarrett, "The Relation of Past Managerial Achievement to Test Measures Obtained in the Employment Situation: Methodology and Results-II", in Personnel Psychology, Vol. 28, 1975, p. 215 - 231.

Overall, the work of Grimsley and Jarrett appear to reflect the general trend noted by Campbell et al. with regard to the potentially valuable role to be played by various psychometric tests in the prediction of managerial progress and effectiveness. This would appear to be the case particularly when the research emphasis rests with a multi-test, multi-dimensional approach involving the observation of patterns of predictor-criterion relationships, and the incremental validity stemming from the "judicious" combination of various paper-and-pencil test scores.

In looking across the various multiple psychometric assessment programs, it is difficult to isolate or determine those dimensions of human functioning and performance which appear to emerge as consistent predictors of managerial effectiveness. Undoubtedly, in part, this is due to the wide variety of test instruments utilized. For example, in Korman's review⁴⁶ of nineteen studies related to the predictive validity of cognitive ability factors, more than twenty-five different test instruments were employed as predictors. As well, particularly in the area of personality testing, different trait names are frequently used to denote the same quality, or the same name is used to represent different qualities.⁴⁷

46 Korman, Op. Cit., 1968, p. 295 - 322.

47 Ghiselli, Op. Cit., 1973, p. 464.

However, within each series of investigations where the focus is on the observation of patterns of significant correlations, certain trends do appear to emerge. For example, Laurent concluded that the more successful managers tended to have higher cognitive ability, better managerial judgment, and were "...more objective, more stable emotionally, and more able in their personal relations."⁴⁸ The results of the investigations of Grimsley and Jarrett^{49,50} suggest that higher level managers were consistently higher on various cognitive abilities as well as showing an orientation suggesting greater activity, vitality, ascendancy and personal relations. Other aspects of personality, interest and values were also found to differentiate the two groups of managers but with less replicability.

In their review of several such research investigations, Campbell et al. point to "...the significant role of dominance, self-confidence, intelligence, and interest in verbal, persuasive, and directive activities for predicting management effectiveness."⁵¹

48 Laurent, Op. Cit., 1968, p. 14.

49 Grimsley and Jarrett, Op. Cit., 1973, p. 31 - 48.

50 Grimsley and Jarrett, Op. Cit., 1975, p. 215 - 231.

51 Campbell et al., Op. Cit., 1970, p. 192.

As an overall concluding statement, they remark:

The construct of effective executiveship...includes such factors as high intelligence, good verbal skills, effective judgment in handling managerial situations..., and organizing skill; dispositions toward interpersonal relationships, hard work, being active, and taking risks; and temperamental qualities such as dominance, confidence, straightforwardness, low anxiety, and autonomy. Moreover, men rating high in overall success report backgrounds suggesting a kind of "life-style" of success....⁵²

2. Assessment Centre Prediction of Managerial Success

Irrespective of the role of situational versus psychometric paper-and-pencil tests, the overall validation research of the assessment centre approach as a means of identifying managerial talent has been highly consistent and impressive. This is perhaps best exemplified by the research into the first major use of the assessment process within an industrial organizational setting.^{53,54,55,56} The

⁵² Ibid., p. 195 - 196.

⁵³ Douglas W. Bray, Richard J. Campbell and Donald L. Grant, Formative Years in Business: A Long-term AT&T Study of Managerial Lives, New York, Wiley, 1974, XVIII-235 p.

⁵⁴ Bray and Grant, Op. Cit., 1966, p. 1 - 27.

⁵⁵ Donald L. Grant and Douglas W. Bray, "Contributions of the Interview to Assessment of Management Potential", in Journal of Applied Psychology, Vol. 53, 1969, p. 23 - 34.

⁵⁶ Donald L. Grant, Walter Katkovsky, and Douglas W. Bray, "Contributions of Projective Techniques to Assessment of Management Potential", in Journal of Applied Psychology, Vol. 51, 1967, p. 226 - 232.

study which was conducted within the American Telephone & Telegraph organization (AT&T) consisted of a sample of 422 men from lower management levels or non-managerial positions. These managerial candidates were assessed over a four year period in the late fifties, with 203 of the total sample being followed up approximately six to eight years later. An important component of this predictive study was that the research data were kept completely confidential and never utilized as part of any selection process thus preventing any possibility of criterion contamination.

The assessment techniques employed were exceedingly diverse and included: an extensive interview, three projective tests, two personality inventories, various cognitive ability tests, an administrative In-Basket simulation, and two group simulations. The test information and behavioural observations gathered through these techniques were then integrated by the assessment staff, leading to an evaluation of performance on twenty-five dimensions deemed to be important for effective managerial performance. As well, there was an overall evaluation of potential centered on the likelihood of each candidate to attain a middle management position in ten years or less. In presenting data on the general predictive validity of this overall assessment centre evaluation of potential, Bray and Grant state:

Approximately 80% of those who had advanced to middle management were judged by the assessment staff as having such potential. The predictions were even more

accurate for those who have not advanced beyond the first level. Most of these men (95%) were judged as lacking in advancement potential.⁵⁷

The results of this and several other researches of the overall predictive validity of the assessment centre approach have been extensively summarized in a major review by Cohen, Moses, and Byham.⁵⁸ These authors evaluated the published research available from 1964 through 1972 in a variety of industrial-organizational settings. Of the nineteen investigations reviewed, all but one presented data which indicated that the overall assessment centre evaluation of managerial potential was predictive of eventual managerial progress or performance. A median predictor-criterion coefficient of .37 was obtained across the various studies. The one qualification to this finding was that the magnitude of the correlation varied depending upon the global criterion measure employed, with a greater relationship being found in predicting job potential (median correlation of .63) than in predicting actual job performance or present job progress (median correlations of .33 and .40, respectively).

The validation research of the assessment centre approach which has been published since the time of this review continues to offer further empirical support for the predictive relationship existing between the overall assessment centre judgment of managerial potential

57 Bray and Grant, Op. Cit., 1966, p. 23.

58 Barry M. Cohen, Joseph L. Moses, and William C. Byham, The Validity of Assessment Centres: a Literature Review, Monograph II, Development Dimensions Inc., Pittsburg, 1974, 1 - 31 p.

and various concurrent and long-term criterion measure of managerial success. In one such investigation⁵⁹ a concurrent validity strategy was used to evaluate the performance of 48 first-line supervisors in a one-day assessment centre. The assessment tools included a management questionnaire, an interview, an In-Basket work sample, a film case discussion, a selection simulation and a management decision game. The overall assessment centre rating of performance for these supervisors was found to correlate at the .47 level with an on-the-job rating of demonstrated and potential performance.

Two additional studies have been reported within the AT&T organization. In the first⁶⁰, over forty-eight hundred non-management women were assessed over an eight year period with follow-up criterion data being gathered two to ten years later. The two-day assessment centre consisted of a variety of management simulations (leaderless group discussion, business game, and In-Basket test) as well as an interview and a variety of paper-and-pencil tests and exercises, including measures of cognitive ability. In using present managerial level as a criterion, Moses and Boehm found that the assessment centre judgment of potential resulted in a correlation coefficient of .37.

59 G.M. Worbois, "Validation of Externally Developed Assessment Procedures for Identification of Supervisory Potential", in Personnel Psychology, Vol. 28, 1975, p. 77 - 91.

60 Joseph L. Moses and Virginia R. Boehm, "Relationship of Assessment Center Performance to Management Progress of Women", in Journal of Applied Psychology, Vol. 60, 1975, p. 527 - 529.

In the second study by Huck and Bray,⁶¹ two samples of non-management women were evaluated in a similar assessment centre. All of these candidates were later promoted to supervisory management positions. The two criterion ratings obtained approximately one to six years later consisted of an overall evaluation of on-the-job performance and the potential for further advancement. With regard to job performance the predictor-criterion coefficients for the two samples were .35 and .41. In terms of potential for further advancement, coefficients of .54 and .59 were obtained.

Mitchel⁶² followed up, one to five years later, 254 managers who were generally at lower management levels at the time of assessment. Once again a variety of situational management exercises were used along with paper-and-pencil tests. For purposes of the study three samples were generated with the predictor-criterion relationship being assessed at three different points in time, that is, one, three, or five years following assessment. Overall, the correlations between the assessment centre evaluation of potential and a criterion measure of salary growth ranged from .10 to .32 with a median coefficient of .22.

61 James R. Huck and Douglas W. Bray, "Management Assessment Centre Evaluations and Subsequent Job Performance of White and Black Females", in Personnel Psychology, Vol. 29, 1976, p. 13 - 30.

62 James O. Mitchel, "Assessment Centre Validity: A Longitudinal Study", in Journal of Applied Psychology, Vol. 60, 1975, p. 573 - 579.

Finally, mention can be made of research by the present author⁶³ related to the assessment of potential for effective functioning at the first level of supervisory management. A total of 117 candidates were evaluated in a two and one-half day assessment centre utilizing techniques which included an interview, five simulation exercises, and paper-and-pencil measures of cognitive abilities, creativity, supervisory judgment, and personality. As with the above-mentioned study, criterion information was gathered at different points in time ranging from one and one-half to four and one-half years following assessment. As well, different criterion measures were employed reflective of both career progression (salary growth and management level) and job performance (two separate ratings of on-the-job performance). Across the two samples generated for this research and across the different time spans, the range of predictor-criterion correlations for measures of career progression extended from .23 to .50 with a median coefficient of .38. With regard to criterion measures of job performance, the range was from .05 to .43 with a median of .26. When the various criteria were combined into an overall composite index of success, the resulting correlations obtained for the two samples were .30 and .50.

63 L.W. Slivinski, K.W. Grant, R.P. Bourgeois, L.D. Pederson and J.L. McCloskey, "Longitudinal Follow-up of a First Level Management Assessment Centre", in Journal of Assessment Centre Technology, Vol. 1, 1978, p. 6 - 11.

In general, the data presented in these five most recent investigations of the assessment centre approach tend to further support the conclusions derived from the review of Cohen et al..⁶⁴ It would appear that little published evidence exists which does not support the existence of a predictive relationship between the global assessment centre judgment and various criterion measures of managerial effectiveness. The magnitude or strength of the relationship does vary considerably from sample to sample and from study to study dependent upon several factors, including the criterion measure or measures employed. In utilizing median correlation coefficients as indicative of the general trend, it would seem that the assessment centre approach can account for 1 to 40 percent of the variance associated with the prediction of managerial effectiveness.

3. Situational Versus Paper-and-Pencil Components of an Assessment Centre

The consistency with which the Assessment Centre approach has been found to be predictive of managerial and organizational success has explicitly or implicitly been used as a means of supporting the value of the situational method of testing over the paper-and-pencil procedure, as well as over traditional management tools such as the selection interview and the assessment of past and present job performance. The rationale appears to be that since the situational

64 Cohen et al., Op. Cit., 1974, 1 - 31 p.

tests represent a major and unique component of the process, validation of the overall assessment centre approach leads to direct or indirect support for the validity of these situational tests.^{65,66}

Increasingly, this has led to a state of affairs whereby assessment centres have tended to downplay, neglect, or omit the role of various psychometric paper-and-pencil tests, and other evaluation tools, in favour of an almost sole reliance on the situational method of testing. Mackinnon, in his invited address to the American Psychological Association, acknowledged this trend:

In recent years psychometric tests have largely disappeared from assessment centre practice, and when administered, test scores are typically not used in making decisions or recommendations concerning candidates. Rather, they are used as checks upon decisions and recommendations already agreed upon... The relative neglect of psychometric tests in assessment centres can be expected to continue for some time.⁶⁷

Much of the support for the importance of situational tests over other assessment tools within the assessment centre has come from data related to internal validation. In these efforts, attempts are

65 Marvin D. Dunnette, "The Assessment of Managerial Talent", in Paul McReynolds (Ed.) Advances in Psychological Assessment, Palo Alto, Science and Behaviour Books, Vol. II, 1971, p. 95.

66 Ann Howard, "On Assessment of Assessment Centers", in Academy of Management Journal, Vol. 17, 1974, p. 126.

67 Donald W. MacKinnon, "Human Assessment: Perspective and Context for Current Practice", Invited address to American Psychological Association, 1975, p. 15.

made to determine the contribution made by the assessment components to the various managerial dimensions evaluated during the assessment centre process, including the final overall recommendation or rating of potential. For example, it may be recalled that in the initial AT&T studies, the information gathered from the various data sources was utilized to obtain ratings on 25 managerial dimensions. For research purposes, these dimensions were factor analyzed leading to the establishment of nine factors. In the various studies,^{68,69,70} the relative contribution of the different assessment techniques to each of the nine factors was determined. Overall, the results suggested that the situational tests (two group simulations and the In-Basket Test) made the largest contribution to the nine factors. This was followed at a second level by the paper-and-pencil cognitive ability tests and by the projective personality test and the selection interview variables. The paper-and-pencil personality inventories showed the least contribution to the factors.

68 Bray and Grant, Op. Cit., 1966, p. 1 - 27.

69 Grant and Bray, Op. Cit., 1969, p. 23 - 34.

70 Grant et al., Op. Cit., 1967, p. 226 - 232.

However, as stated by MacKinnon, this does not in itself serve to demonstrate the lack of validity of such assessment devices.

The reported failure of such tests in some studies to contribute importantly to the overall final rating or recommendation may be more a sign of the way in which test scores were used or even not used by assessors than an indication of their irrelevance to the making of valid assessment centre predictions. In other words, the fault or lack may be more in the assessors than in the tests themselves.⁷¹

Obviously a more relevant means of evaluating the role of situational and paper-and-pencil components would be through external predictive validation. The difficulty, however, is that much of the predictive validity research related to assessment centres has concentrated on the validity of the overall final judgment made by the assessors, a judgment which is arrived at by combining the information from several data sources including both situational and paper-and-pencil tests. Some investigations^{72,73,74,75} do present individual predictor-criterion correlations for a diversity of situational and

71 MacKinnon, Op. Cit., 1975, p. 16.

72 V. Jon Bentz, "Validity of Sears Assessment Centre Procedures", Paper presented at American Psychological Association Washington, 1971.

73 Bray and Grant, Op. Cit., 1966, p. 1 - 27.

74 Lloyd D. Marquardt, Follow-up Evaluation of the Second Look Approach to the Selection of Management Trainees, Chicago, Sears/Roebuck, 1976.

75 Paul E. Wernimont, Management Assessment: Relationships Between Test Scores and Supervisor's Ratings on Job Performance Dimensions, Research Report No. 67, 3M Company, 1973.

paper-and-pencil measures, but the absence of statistical techniques to compare the contribution of one relative to another prohibits any firm conclusions. Even meaningful trends are difficult to evaluate due to the wealth of correlations presented. For example, one research study⁷⁶ contains 33 pages of predictive validity coefficients, all of them statistically significant; another article⁷⁷ includes consideration of 27 separate criterion measures.

When comparative statistical techniques are employed, the main approach has been to compare the predictive validity of a single assessment component with the validity associated with the overall assessment centre judgment. This approach is subject to interpretive and methodological difficulties since the overall assessment judgment is itself derived from several data sources, including the component being isolated for special analysis.

As one example of this approach, Bray and Campbell⁷⁸ followed up 78 men who had participated in a two-day assessment centre consisting of three situational tests, four cognitive paper-and-pencil tests, a lengthy interview and biographical data. The information generated from

76 Bentz, Op. Cit., 1971.

77 Marquardt, Op. Cit., 1976.

78 Douglas W. Bray and Richard J. Campbell, "Selection of Salesmen by Means of an Assessment Centre", in Journal of Applied Psychology, 1968, Vol. 52, p. 36 - 41.

these various data sources were utilized to evaluate each participant on twenty assessment dimensions, leading to an overall final recommendation of potential. A field rating of overall job performance obtained six months after assessment was employed as the main criterion measure. As part of the results, the authors found that the final assessment recommendation correlated .51 with the performance criterion, and that this was higher than the multiple correlation of .33 obtained with the four paper-and-pencil cognitive ability tests. At best, one might conclude from these results that an assessment program consisting of situational tests, a management interview, biographic information and paper-and-pencil cognitive ability tests resulted in a higher level of predictive validity than that obtained from the abilities tests alone.

Other studies have also employed this type of comparison particularly as a means of assessing the contribution of paper-and-pencil measures of cognitive ability,^{79,80} although some attempts have also been made to evaluate other data sources such as projective tests

79 Bray and Grant, Op. Cit., 1966, p. 1 - 27.

80 Joseph L. Moses, "Assessment Centre Performance and Management Progress", in Studies in Personnel Psychology, Vol. 4, 1972, p. 7 - 12.

and the interview.^{81,82} The consistent pattern which emerges from these studies is that the combination of multiple assessment procedures results in higher predictive validity than does any single assessment procedure. The studies, however, do not directly compare the situational and paper-and-pencil components of an assessment centre, unless one is willing to assume that, since the situational tests seem to make a greater contribution to the overall assessment evaluation, they probably in turn account for the higher degree of predictive validity obtained. As well, the studies do not broach the issue as to whether a "judicious" combination of paper-and-pencil tests (e.g. combining measures of cognition with those of personality) might result in validity coefficients superior to those obtained from the much more expensive and extensive situational testing approach.

Only two studies have employed statistical procedures which allow for a direct comparison of the individual and incremental contribution made by the situational and paper-and-pencil components of an assessment centre. The first is a study reported by Glaser, Schwarz

81 Robert M. Finley, "Evaluation of Behaviour Predictions from Projective Tests Given in a Management Assessment Centre", Paper presented at the American Psychological Association, 1970.

82 F.O. Carleton, "Relationships between Follow-up Evaluations and Information Developed in a Management Assessment Centre", Paper presented at the American Psychological Association, 1970.

and Flanagan in 1958.⁸³ They utilized a concurrent validity strategy with eighty civilian supervisors within a military setting in order to compare scores on two short paper-and-pencil tests with ratings obtained through three management simulation exercises. The psychometric tests consisted of a measure of basic cognitive ability and a test of supervisory judgment. The situational tests involved a group discussion, a small-job management problem, and a personnel problem.

The concurrent criterion measure was a composite index of supervisory performance. The individual correlations of the basic cognitive ability and supervisory practices tests to the criterion were .25 and .23, respectively, with a combined multiple correlation of .27. The individual correlations of the three situational tests were .08, .16, .17, with the small-job management situation showing the least relationship. The multiple correlation combining the three simulations was not presented by the authors. When the situational component was added to the paper-and-pencil component, the multiple correlation coefficient was raised from .27 to .33. These results were generally interpreted by the authors in a positive light. "In general, these findings of the study indicate that some contribution to the predictive value of the paper-and-pencil tests can be made by the

⁸³ Robert Glaser, Paul A. Schwarz, and John C. Flanagan, "The Contribution of Interview and Situational Performance Procedures to the Selection of Supervisory Personnel", in Journal of Applied Psychology, Vol. 42, p. 68 - 73.

addition of the predictors studied."⁸⁴

A more extensive investigation emphasizing an evaluation of the relative and incremental validity of situational and paper-and-pencil tests was conducted by Wollowick and McNamara in 1969.⁸⁵ The ninety-four subjects followed up by these authors were generally from lower to middle management levels at the time of participation in a two-day assessment centre program. The assessment techniques consisted in part of six group and individual situational tests. For each of these simulations, the subjects were evaluated on twelve characteristics considered important for managerial performance. As well, an overall evaluation of performance was given for each of the six situational exercises. The paper-and-pencil component of the assessment process consisted of eight measures of personality dimensions, two measures of leadership style, two indices of general cognitive ability, and two measures emerging from biographical data, for a total of 14 scores.

As expressed by Wollowick and McNamara, they actually evaluated three assessment components: tests, characteristics, and exercises. The "test" component was comprised of the 14 paper-and-pencil scores. The "characteristics" component entailed the average rating given to each of the twelve managerial dimensions across the various exercises.

⁸⁴ Ibid., p. 73.

⁸⁵ Herbert B. Wollowick and W.J. McNamara, "Relationship of the Components of an Assessment Center to Management Success", in Journal of Applied Psychology, Vol. 53, 1969, p. 348 - 352.

The "exercises" component consisted of the six overall ratings given to each of the six management simulations. The criterion measure utilized was an index of the career progression achieved approximately three years subsequent to assessment.

As part of their research, Wollowick and McNamara employed a series of multiple regression analyses. The results of these analyses indicated that all three assessment components of tests, characteristics and exercises were significantly related to the criterion, with multiple correlations of .45, .41, and .39, respectively. The psychometric data accounted for approximately 20% of the variance, the managerial dimensions 17%, and the simulation exercises 15%. When the three components were combined, a multiple correlation of .62 was obtained, accounting for 38% of the variance. The authors conclude that "...all three types of variables contribute heavily to the predictive success of the program".⁸⁶

These results are important in suggesting the role to be played by different assessment methods. One limitation or methodological problem with the study is that there was possible confounding or contamination of the situational test measures. The assessors in determining their situational judgements were not aware of the paper- and pencil test scores. However, in addition to observing performance in the management simulations, the assessors also conducted interviews with the participants. As well, they had access to information related

86 Ibid., p. 351.

to each participant's career history and biographical background.⁸⁷

It is likely that the judgments made as a result of performance in the situational tests were influenced or contaminated to some degree by knowledge obtained from these other data sources. From this perspective, the abovementioned study may perhaps be more appropriately viewed as one in which a paper-and-pencil testing approach was compared with an assessment program consisting of situational tests, interview data and biographic and career-related information.

Perhaps more importantly, however, is the fact that only two studies over the last two decades have made attempts to directly evaluate the role of paper-and-pencil and situational tests, and only one of these studies was predictive in nature. Over the past few years various authors^{88,89,90} who have reviewed the general assessment centre process have commented on the lack of research efforts in this area, and have stressed the need for further experimentation in determining the contribution to be made by different assessment tools, many of which have been traditionally incorporated under the entity known as

87 Herbert B. Wollowick, Personal communication, 1978.

88 Richard J. Klimoski and William J. Struckland, "Assessment Centres: Valid or Merely Prescient", in Personnel Psychology, Vol. 30, 1977, p. 353 - 361.

89 Bernard Ungerson, "Assessment Centres - a Review of Research Findings", in Personnel Review, Vol. 3, 1974, p. 4 - 13.

90 John E. Wilson and Walter A. Tatge, "Assessment Centres: Further Assessment Needed?", in Personnel Psychology, Vol. 26, 1973, p. 172 - 179.

the assessment centre. For example, in one of the first reviews of this type, Wilson and Tatge state:

The critical question is...whether they (assessment centres) contribute significantly more to the validity of the selection decision than paper-and-pencil tests, management assessment of past experience and performance, psychological appraisals, interviews, weighted experience records, or the various combinations of these measures.⁹¹

After examining the available evidence the authors concluded with this statement:

In summary, assessment centres have been established as a valid selection method, but considerable research is needed to justify claims for superiority over the traditional methods.⁹²

Ungerson in a similar review of the literature echoes the same comment:

It is not enough to show that they (assessment centres) are valid... There should be massive experiments...to compare assessment centre predictions with other methods.... Until this kind of research has been carried out on an adequate scale, enthusiasm should be tempered by scientific scepticism.⁹³

In the most recent critical review of the assessment centre process, Klimoski and Struckland offer the following conclusion:

The point we would like to make is that little published research exists comparing alternative predictors with assessment centre predictions...we must entertain the

91 Ibid., p. 173.

92 Ibid., p. 179.

93 Ungerson, Op. Cit., 1974, p. 13.

notion that there are more easily obtained predictors available than the assessment centre, and these alternative predictors might offer superior validities.⁹⁴

In addition to these critiques of the general assessment centre approach, other authors^{95,96,97} have focused more specifically on the role to be played by psychometric paper-and-pencil tests, in contrast to the situational testing component of the assessment process. Howard, in her review of the reserach literature, suggests that, although the assessment centre looks promising, "...the psychometric literature...has not been disproved".⁹⁸ At another point, she concludes that "...more research evidently is needed on the various components of assessment centres and their integration..."⁹⁹

Huck, in a similar review of the research, concluded by stating that "...numerous questions regarding the assessment process remain unanswered...and a critical need exists for additional research dealing with the assessment centre approach..."¹⁰⁰ As one of the

94 Klimoski and Struckland, Op. Cit., 1977, p. 357.

95 Howard, Op. Cit., 1974, p. 115 - 134 .

96 James R. Huck, "Assessment Centers: A Review of the External and Internal Validities", in Personnel Psychology, Vol. 26, 1973, p. 191 - 212 .

97 MacKinnon, Op. Cit., 1975, p. 1 - 31.

98 Howard, Op. Cit., 1974, p. 131.

99 Ibid., p. 126.

100 Huck, Op. Cit., 1973, p. 209.

research directions to be taken, Huck acknowledges that "...multiple assessment procedures provide a number of data sources, and the contribution of each to...multiple criterion measures must be clarified...".¹⁰¹

Finally, in his invited address to the American Psychological Association, MacKinnon strongly urged that more researches be conducted into the role of psychometric paper-and-pencil testing. In part, he states:

Until they have been carried out we shall not know whether or to what degree and in combination with what other variables various psychometric tests can indeed contribute to valid prediction of criteria of managerial effectiveness.¹⁰²

4. Summary and Statement of the Basic Hypotheses

Research on the assessment centre process indicates that this multi-method evaluative approach can be utilized as an effective tool in the identification, selection, and prediction of managerial talent. However, few controlled and systematic investigations have been carried out to determine the relative and incremental predictive validity associated with the various testing methods encompassed within the assessment centre process. The focus of the present research is to compare two such major components of the assessment centre process, namely, the situational testing method and the psychometric paper-and-pencil approach to testing.

¹⁰¹ Ibid., p. 209.

¹⁰² MacKinnon, Op. Cit., 1975, p. 18.

A general review of the psychometric literature suggests that various individual dimensions of behaviour, assessed through paper-and-pencil tests, result in at best low to moderate degrees of predictive validity. However, when a multiple predictor approach is taken through the employment of a variety of paper-and-pencil tests or dimensions, a higher degree of predictive validity appears to ensue, with one estimate suggesting that as much as 30 to 50 percent of the variance associated with effective managerial performance being accounted for through this testing approach.

Much of the support in favour of the situational approach to testing has been established indirectly through the validation research of the assessment centre method, of which the situational component appears to be a prime contributor. Overall, the research in this area consistently illustrates the predictive validity of the assessment centre, with as much as 40 percent of the total predictor-criterion variance being taken into account in some investigations.

Studies which have directly compared the situational and paper-and-pencil components of an assessment centre have been few in number, and there is an important need for further research in this area. In the present study, an attempt is made to compare the predictive validity of a battery of paper-and-pencil tests, consisting of measures of intelligence, personality and administrative judgment, with a series of five individual and group situational tests.

Although this research did not deal directly with the theoretical issues surrounding assessment the results can easily be used to shed some light on some of the themes inherent in assessment. One could use the following framework to help bring into focus these various themes. Along one dimension there is the sign-sample data controversy. It is exemplified by the instrument used to gather information about the individual. At one end of the continuum are paper-and-pencil tests (sign data) and at the other is situational testing (sample data). Between these two poles lies biographical information, interviews, case methods, incident method and a number of other measurement techniques. The second dimension concerns itself with the criterion. Global, singular, multiple and composite criteria make up this continuum. The third dimension includes the methodology of data-gathering/data-combination which can be clinical/judgemental or actuarial/mechanical. With this distinction Sawyer^{102a} arrives at six prediction strategies.

Overlaid on these three dimensions are two templates, so to speak. As mentioned previously, research has shown that validity is time-locked. Therefore, validity must be established at varying points in time. The other template has to do with the validation model. Basically this concerns itself with subgrouping according to tests, people, work behaviours, and organizational outcomes.

102a J. Sawyer, "Measurement and Prediction, Clinical and Statistical", in Psychological Bulletin, Vol. 66, 1966, 178-200.

REVIEW OF THE LITERATURE AND
STATEMENT OF THE PROBLEM

37b

This thesis when applied to the above model can be classed as one where sign and sample techniques are used to gather clinical and actuarial data combined in an actuarial mode against global and composite criteria at one point in time with no subgrouping.

The first major hypothesis expressed in its null form reads:

There is no significant difference in the amount of variance accounted for by a set of paper-and-pencil test measures and a set of situational test measures in the prediction of managerial effectiveness.

An additional major concern in this research will be not only to determine the predictive validity of one component relative to another, but also to establish the amount of incremental validity which ensues when one component is combined with the other.

Thus, two additional hypotheses are generated which, when stated in their null form, read as follows:

There is no significant difference in the incremental validity added by a set of paper-and-pencil test measures when the validity associated with a set of situational test measures is taken into account.

There is no significant difference in the incremental validity added by a set of situational test measures when the validity associated with a set of paper-and-pencil test measures is taken into account.

CHAPTER 2

EXPERIMENTAL DESIGN

This chapter details the methodology related to the present research investigation. The initial section describes the subjects participating in the study. A second section places the research within the context of the assessment centre program for the Federal Public Service. This is followed by a description of the paper-and-pencil and situational test measures, along with a delineation of the criterion measures utilized. The fourth section details the relevant testing and scoring procedures. The concluding section describes the statistical procedures employed to test the hypotheses noted in the first chapter.

1. Subjects

The Ss for this study were participants in any one of seven assessment centre programs held between March, 1973 and May, 1975. Of the total population of 322 participants, it was possible to obtain complete predictor and criterion data on a sample of 297. This represents 92% of the total group. All of the Ss were Federal Government of Canada employees who were generally at lower middle to upper middle levels of management at the time of assessment. Forty six different government departments, agencies, and crown corporations were represented in the sample. Eighty percent of the Ss were assessed in the English language, with the remainder being evaluated in French. Eighty-four percent of the participants were male.

With respect to age, the range extended from 26 to 49 years, with an average of 35 years. Sixty-five percent of the Ss were in their thirties at the time of assessment. In terms of educational background, ninety-three percent possessed some type of university or college degree or diploma. Thirty-three percent of the sample held a post-graduate degree.

In terms of occupational status, the Federal Public Service differentiates two prime occupational categories under which are subsumed various classification groups. Seventy-five percent of the Ss in this study were from the Administrative category, with the remainder being from the Scientific and Professional category. Within these two broad categories, twenty-eight different classification groups were represented. Major representation, accounting for 50% of the sample, came from three main classification groups known as Program Management, Personnel Specialists, and Administrative Services. The general career progression of the Ss up to the time of assessment was quite varied. On average, the participants had worked for a total of approximately thirteen years, of which approximately eight years had been spent within the Federal Public Service. However, with respect to these two variables, the range extended from three to thirty years and from one to twenty-five years, respectively. As well, participants had spent an average of two and one-half years working in their most recent position prior to participation in the assessment centre. Here, the range extended from one to twenty two years. Appendix I contains frequency distributions and related statistics for the subject variables noted above.

2. The Context of the Assessment Centre Program

In order to place in perspective the measuring instruments employed in the present investigation, this section describes the context of the assessment centre program as it is used within the Canadian Federal Public Service. The assessment centre program is used as part of the general selection process for the federally sponsored Career Assignment Program (CAP). The CAP program is an integrated executive education and development program designed for middle level managers who have demonstrated high potential to eventually assume executive level responsibilities within the government. Initial nomination to the CAP program is determined by the sponsoring department according to their own screening procedures but based on a series of guidelines. These include requirements that the nominee be employed by the sponsoring organization for at least two years, have a superior record of performance with demonstrated above-average potential for advancement to executive levels, and be generally at the middle level of management in terms of salary and/or responsibility.

As part of the overall selection process, all nominees participate in the two and one-half day assessment centre program. The prime objective established for the assessment centre is to assist departments in their selection of CAP participants by providing a standardized testing system through which further information can be gathered regarding the executive potential of nominees. To achieve this the assessment centre program employs two assessment methodologies, namely, situational tests and paper-and-pencil measures of performance.

Unlike most other assessment centre programs, other evaluative tools such as the interview and the use of biographical and career history information is not incorporated into the assessment centre process.

Throughout the development of the assessment centre, the focus centered on the utilization of evaluative tools which would assist in the identification and development of executive potential. The development of the situational tests and the selection of paper-and-pencil tests emerged through extensive research of executive positions within the Federal Public Service. In part, these investigations included the use of diary records and structured interviews designed to establish the common executive duties and responsibilities which cut across departments, organizational units and occupational specialties; job analysis questionnaires in order to determine the various job functions and job types associated with executive management; survey instruments designed to establish the important and relevant executive qualities or performance dimensions. The simulation exercises were developed as vehicles to elicit the qualities established as important for effective performance as an executive within the Federal Government. The background development and implementation of the assessment centre program is extensively detailed in technical report form.¹⁰³

The end result of the assessment centre process is the generation of an overall assessment report. The contents of the report are fed back to each nominee by a psychologist within a career counseling framework. It is the nominee who decides whether to proceed

¹⁰³ L.W. Slivinski, R.P. Bourgeois, L.P. Pederson and J.L. McCloskey, Development and Application of the Assessment Centre for the identification of Senior Executive Potential, Public Service Commission of Canada, In Press, 1978.

further with the nomination to the CAP program. To do so requires a release of the results of the assessment report to a committee within the sponsoring organization whereby the members are able to read but not retain the report. No information regarding performance at the assessment centre is retained on file within a department.

The final decision regarding acceptance onto the CAP program is determined by a selection review board consisting of senior officials from selected government departments. This decision is based on the information contained in the assessment report and on a variety of departmental inputs including the nominee's general career background and progression, his/her identified performance record on the job, and the developmental recommendations of the sponsoring department. On occasion, a personal interview may be held with some nominees by the selection review board.

3. The Measuring Instruments

a. Paper-and-Pencil Test Measures:

The paper-and-pencil test battery was selected to measure the three areas of cognitive ability, administrative judgment, and personality. These areas represent three aspects of behaviour which the review of management literature showed to hold value as potential predictors of success within organizations. The tests which comprised the paper-and-pencil battery are described below.

i) Cognitive Ability: A measure of general intellectual functioning was determined for this study by means of the General

Intelligence Test (GIT).¹⁰⁴ This test was developed by the Public Service Commission of Canada for use for personnel selection and developmental purposes within the federal government. In adopting the Wechsler Adult Intelligence Scale as a working model, the GIT was designed to test a broad range of intellectual abilities normally associated with one's general adaptiveness across diverse tasks and situations. The test consists of six subtests and leads to an overall measure of general intellectual ability. A description of the various subtests which comprise the GIT, along with representative item examples, are noted in Appendix II. With respect to test reliability, the manual¹⁰⁵ details various Kuder-Richardson 20 internal consistency estimates of reliability. Overall, the internal consistency for the GIT is found to be acceptably high with coefficients ranging from .92 to .94.

ii) Administrative Judgment: The research efforts of Laurent¹⁰⁶ demonstrated the consistency with which a paper-and-pencil measure of managerial judgment could be related in a predictive fashion with managerial level. Unpublished data by the present author has also suggested the value of a supervisory judgment test in predicting pro-

¹⁰⁴ General Intelligence Test Manual (Second Edition), Personnel Psychology Centre, Public Service Commission of Canada, 1977, 1 - 40 p.

¹⁰⁵ Ibid., p.22.

¹⁰⁶ Laurent, Op. Cit., 1970, p. 417 - 423.

gression to first-level managerial responsibilities.¹⁰⁷ In the present investigation, a measure of a manager's broad understanding of the process of administration of large organizations was evaluated by means of a 50-item multiple choice questionnaire known as the Administrative Judgement Test.¹⁰⁸ Originally employed within the U.S. Civil Service, the test was adopted for use within the context of the Canadian Federal Government.

The items of the Administrative Judgement Test endeavour to measure the common elements in the administrative process. They include problems in the relationships between the headquarters and field offices, between research and operating personnel, as well as problems on the timing of programs and the organization of the office of an administrator. The test was not designed to measure technical knowledge. Detailed in the manual¹⁰⁹ are various reliability scores evaluated by means of Kuder-Richardson 20 formula, as well as the split-half estimate of reliability. Across various samples, the reliability coefficients ranged from .65 to .94.

iii) Personality: Although the research reviewed in the introductory chapter pointed out the possible predictive power of paper-and-pencil personality measures, wide variation in the dimensions explored,

107 L.W. Slivinski, "The Relation of the Supervisory Judgement Test to Career Progression and Job Performance", Unpublished Data, 1977.

108 Administrative Judgement Test Manual, U.S. Civil Service Commission, Technical Series, No. 34, 1966, 1 - 15 p.

109 Ibid., p. 9.

and the trait labels given, was in evidence. For the present investigation, seven personality components of the Personality Research Form (PRF)¹¹⁰ were isolated for consideration. The PRF actually evaluates twenty-two different dimensions of personality. In selecting the seven dimensions of personality for investigation, the author was guided by the "superordinate" categories established through factor analytic studies of the PRF scales.¹¹¹ The manual delineates seven broad categories under which the twenty-two scales can be subsumed. These represent measures of orientation toward work and play; degree of ascendancy; orientation towards direction from other people; degree and quality of interpersonal orientation; intellectual and aesthetic orientation; impulse expression and control; and test-taking attitudes and validity. Of these, the first four categories appear to offer the greatest linkage with the patterns of personality behaviour observed in various management studies.

The prime measures related to an individual's work orientation on the PRF involve the dimensions of "achievement" and "endurance". The defining trait adjectives for these scales reflect in general similar qualities noted in other research investigations, including such aspects as hard work, general activity, ambition, energy, vitality, and

110 Douglas N. Jackson, Personality Research Form Manual, Research Psychologists Press, Goshen, 1967, 1 - 31 p.

111 Ibid., p. 4 - 5.

competition. A description of high scorers on these two dimensions would be as follows:

Achievement: Aspires to accomplish difficult tasks; maintains high standards and is willing to work toward distant goals; responds positively to competition; willing to put forth effort to attain excellence.

Endurance: Willing to work long hours; doesn't give up quickly on a problem; persevering, even in the face of great difficulty; patient and unrelenting in his work habits.

With respect to the superordinate category of degree of ascendancy, the two relevant dimensions are "dominance" and "abasement". Many related trait labels have been noted to emerge in the validity research of managerial progress and performance: ascendancy, dominance, leadership, assertiveness and self-confidence.

Dominance: Attempts to control his environment, and to influence or direct other people; expresses opinions forcefully; enjoys the role of leader and may assume it spontaneously.

Abasement: Shows a high degree of humility; accepts blame and criticism even when not deserved; exposes himself to situations where he is in an inferior position; tends to be self-effacing.

The third category involves measures of one's orientation towards direction from other people, and here the defining trait adjectives would appear to center on the concept of independence or self-reliance. The two PRF scales subsumed under this category are "autonomy" and "succourance".

Autonomy: Tries to break away from restraints, confinement, or restrictions of any kind; enjoys being unattached, free, not tied to people, places or obligations; may be rebellious when faced with restraints.

Succourance: Frequently seeks the sympathy, protection, love, advice, and reassurance of other people; may feel insecure or helpless without such support; confides difficulties readily to a receptive person.

The fourth category subsumes measures of the degree and quality of one's interpersonal orientation. This would appear to offer a close linkage with the research pattern suggesting the value of a general personal relations element in the prediction of management effectiveness. The PRF lists several dimensions in this category, with the most general overall estimate of one's interpersonal orientation being that of "affiliation".

Affiliation: Enjoys being with friends and people in general; accepts people readily; makes efforts to win friendships and maintain associations with people.

A more extensive description of all of the PRF dimensions is given in Appendix III. In terms of reliability, the manual¹¹² summarizes various investigations of the Kuder-Richardson formula 20 of internal consistency, as well as test-retest and parallel form reliability estimates. For the seven dimensions used in the present

112 Ibid., p. 20 - 22.

study, the range of reliability coefficients obtained through these three methods extended from .86 to .92, from .75 to .88, and from .66 to .86, respectively.

b. Situational Test Measures:

The situational test measures were evaluated through the observation of performance in five simulation exercises designed to reflect the duties and responsibilities of executive functioning within the Federal Public Service. The general nature of the exercises are highly comparable to those employed in most major assessment centre programs. For example, one emphasizes the evaluation of administrative skills throughout an in-basket exercise. Another centers on an individual case presentation, while a third focuses on a relatively unstructured group discussion.

One major divergence from most traditional assessment centre programs is that all of the simulation exercises are interrelated and integrated in order to reflect the general context of executive management. Each participant, in coming to the assessment centre, immediately assumes the role of an executive in charge of a particular organizational unit within a newly formed government agency. In this role, the participant then proceeds to engage, over a two and one-half day period, in a series of duties and meetings related to the unit. This begins with work on the in-basket which has accumulated (In-Basket

Test) followed by a meeting with the executive's subordinate staff (Staff Meeting). An opportunity then arises to meet the president of the agency (Meeting with the President), followed by a meeting with representatives of another government agency (Treasury Board Presentation). Finally, the candidate is requested to participate in a group discussion related to the problems being faced by both the organizational unit and the agency as a whole (Task Force Exercise). A brief description of each situational test follows.

In-Basket Test: In this initial simulation, the participant is presented with an in-basket in which letters, reports, memoranda, messages, background material, and other papers are deposited for the manager's attention and action. The content represents a judicious choice of paper work representing a cross-section of the range of problems and responsibilities that an executive may encounter on a day-to-day basis. The participant is asked to work on the problems and make decisions within a predetermined time frame (two and one-half hours), recording all responses, decisions, explanations, thoughts, etc. on paper. Certain supporting documentation is left with the participant for use in the forthcoming simulations.

Staff Meeting: This one hour meeting provides an opportunity for the participant to meet four subordinate staff (portrayed by trained actors) in order to question them concerning the state of the organizational unit in preparation for a briefing with the president of the agency. The staff members make individual presentations to their new "boss" reinforcing certain issues previously noted in the in-basket, as well as raising new ones. In this assigned leadership role, the candidate is instructed that he/she may question, comment, direct, or enquire about any material presented by the staff or the in-basket.

Meeting with the President: Following the staff meeting, the participant is given one hour of preparation time prior to meeting and briefing the president. This fifty minute meeting with the president is relatively unstructured and provides an opportunity to observe the participant in interaction with a superior (the role being portrayed by an actor), bringing to bear all relevant material obtained from the two previous simulations. The format and content of the meeting is

determined by the participant, although the president is able to question and converse on all issues.

Treasury Board Presentation: For this simulation, the participant is given the task of analyzing a set of budget figures using back-up materials provided by the subordinate staff, in preparation for a meeting with two representatives of Treasury Board. In preparing for this forty minute oral presentation, the participant must synthesize a considerable amount of quantitative information related to the budget forecast and plans of the organizational unit for the next three years. The participant must also be prepared to respond to a variety of budget-related questions to be posed by the external representatives.

Task Force Exercise: In the final situation, the participant becomes part of a task force established to study and recommend solutions to problems within the organizational unit as well as the agency as a whole. The task force is comprised of six participants, all of whom have gone through the same situational tests. An hour of individual preparation is followed by two hours of group discussion. In this unstructured, leaderless situation, the members are expected to organize time and resources, to attempt consensus on plans of action, and to prepare a summary report for the president.

The five simulation exercises produce two sources of situational test measures. The written responses in the In-Basket Test are evaluated, resulting in ratings on six administratively-oriented dimensions. Performance across the other four simulations is observed, noted and eventually integrated, leading to ratings on eight managerial dimensions.

With respect to the In-Basket simulation, the following dimensions are relevant to the present investigation:

Initiating Activities: This dimension reflects the number of activities initiated in preparing to make decisions such as discussing with others; asking for further information, advice, opinion, or investigation; setting up new procedures or routines; setting up agendas or schedules of "things to do"; replying to others in writing as an interim action; and other actions which imply further involvement of the participant.

Decision Making: This dimension reflects the decisions made by the participant in terms of giving directions and suggestions; concurring or not with the suggestions, recommendations, or requests of others; planning follow-up of situations or delegated work; communicating critical awareness of the quality of suggestions, actions, or performance of subordinates; and making decisions on whether to take action. The appropriateness or quality of the decisions is taken into account.

Problem Comprehension: This dimension reflects the quantity and quality of the participant's comprehension of the issues as evidenced by the consideration given to information from background material and other items; implication of issues not inherent or mentioned in an item; priority or urgency of actions; policy or precedents; and evidence of underlying problems in a situation and the need for alternative approaches.

Delegation: This dimension reflects the number of times problems are referred to subordinates for action. Included in this dimension is the appropriateness of the delegation.

Planning and Organizing: A score on this dimension reflects the number of times work is scheduled for a definite time, and the extent to which items have been grouped for more efficient handling and action.

Productivity: This final dimension reflects overall productivity as evidenced by the number of items handled, as well as the number of courses of action taken.

Across the other four simulations extending from the Staff Meeting through to the Task Force Exercise, the following eight dimensions are evaluated.

Stress Tolerance: The ability to respond rationally, keep control and adapt planned behaviour to unexpected interruptions, diversions or situations.

Motivation: Shows active interest and involvement in tasks presented. Appears to be highly committed with a need to excel. This dimension also includes such aspects as achievement motivation, strong desire to work, ambition.

Independence (of thought and action): A willingness to state one's own personal convictions and to defend them under pressure and opposition; a willingness to identify and pursue courses of action on one's own; a willingness to stand alone,

to depend on oneself; having little need for clear assignment and direction; showing a considerable degree of toughness and courage, having little need for help, advice, direction and encouragement.

Leadership: The ability to influence, convince, direct and persuade others in order to get the job done. Takes the lead in a group and appears to be able to maintain it by controlling; by quality or quantity of contributions; by commenting positively or negatively to contributions of others; by charismatic qualities; or a combination thereof.

Interpersonal Relations and Awareness: The ability to get along with others; to show flexibility in dealing with a variety of personality types; to be tactful, open and fair in dealing with staff; ability to meet people readily, to interact warmly; ability to interact or respond appropriately to the "personal" situation at hand.

Analysis and Synthesis: The ability to detect conditions which lead to potential problems or successes; sensitivity to problems; the ability to analyze problems into their component parts and propose alternative solutions; the capacity to put together diverse ideas or information from a variety of sources to arrive at a solution or plan of action.

Oral Communications: The ability to transmit ideas and information with clarity, a good choice of words and poise; to be interesting, articulate and easy to understand; to use good wording, grammar, syntax and semantics.

Quality of Judgement: Ability to reach sound decisions based on the evidence at hand concerning specific problems and general approaches to people and situations. Does not allow one's own personal biases and prejudices to distort the facts or problems.

In terms of reliability estimates, a recent review by Prien¹¹³ summarizes the reliability research on various in-basket test exercises. With regard to internal consistency measures of the in-basket technique, median reliabilities were found to be generally in the range of .40 to

¹¹³ Erich Prien, "The In-Basket Test: A Review of Research and Opinion", Unpublished Document, 1977, p. 1 - 69.

.60. With respect to the specific in-basket dimensions used in this study, research by the present author¹¹⁴ has shown inter-rater reliability estimates extending from .83 to .93 for an in-basket test which served as a precursor to the present in-basket. For the eight dimensional ratings evaluated across the other four situational tests, unpublished data on samples similar to the one used in the present investigation have resulted in inter-rater reliability scores ranging from .76 to .86.¹¹⁵

c. Criterion Measures:

The vast majority of studies reviewed in the first chapter have employed various singular global criterion measures reflective of the general career progression achieved following assessment or the overall job performance of the individual. For the purpose of the present predictive validity investigation, two global indices of an individual's progression up the managerial or organizational hierarchy were isolated for analysis, namely, position-level and salary. A third criterion measure was arrived at by combining these two criteria into a composite index of "success" within an organization.

i) Position-Level: This criterion measure was established to reflect the general organizational level attained by each assessment centre participant at the time of follow-up. As has been noted in the description of the sample, the Ss involved in this predictive validity

114 R.P. Bourgeois and L.W. Slivinski, "The Inter-rater Reliability of the Consolidated Fund In-Basket", in Studies in Personnel Psychology, Vol. 6, 1974, p. 47 - 52.

115 Vince McDonald and S. Ricard, "Intra-class Correlations Among Assessors' Integration Ratings", Unpublished Document, 1977, p. 1 -6.

study came from any one of twenty eight different classification groups. To this end, a modified version of a standard equivalency table developed by the Federal Public Service was employed. Such an equivalency table, which is noted in Appendix IV, allows for the comparison of an individual within a specific classification group and level with individuals in any other classification group.

In employing this criterion measure consideration was also given to the fact that varying organizational levels were represented at the time of assessment. As well, because the assessment centre program extended over a two year period, the follow-up time interval also varied among participants. These two factors were taken into account by expressing the criterion measure as a rate of increase in position-level. This measure was determined by means of the following equation.

$$\begin{array}{l} \text{Rate of} \\ \text{Position-Level} \\ \text{Increase} \end{array} = \frac{\text{Position-Level at} \\ \text{Time of Follow-up} - \text{Position-Level at} \\ \text{Time of Assessment}}{\text{Assessment Follow-up Time Interval}}$$

ii) Salary: Similar to the position-level criterion, this global measure of the general progression achieved by the assessment centre participants was expressed as a rate of salary increase measure.

$$\begin{array}{l} \text{Rate of} \\ \text{Salary} \\ \text{Increase} \end{array} = \frac{\text{Salary at Time} \\ \text{of Follow-up} - \text{Salary at Time} \\ \text{of Assessment}}{\text{Assessment Follow-up Time Interval}}$$

One difficulty associated with the use of a salary criterion is the fact that increases in salary over time may reflect more general inflationary factors than any significant progression within the organization. In an effort to control for inflationary influences, the amount of salary increase which could be attributed to collective bargaining agreements within the federal government was determined for each S. The salary data used in this study therefore represents the salary increase attained above and beyond that which could be attributed to the normal increases resulting from collective bargaining agreements.

iii) Composite: Preliminary analysis of the above two criterion measures indicated that they were intercorrelated to a moderate degree (.41). Research by Sluckin^{116,117} has indicated that, in addition to exploring each criterion variable individually, measures showing moderate degrees of intercorrelation are ideally suited to combining into a composite index. It is his contention that there is little value in combining highly intercorrelated criterion variables since one is largely redundant of the other. Similarly, criterion measures showing low intercorrelations suggest a distinctiveness not conducive to combining into a composite. In establishing this third criterion measure, the salary and position-level data were converted to a common

116 W. Sluckin, "Combining Criteria of Occupational Success: Part I", in Occupational Psychology, 1956, Vol. 30, p. 20 - 26.

117 W. Sluckin, "Combining Criteria of Occupational Success: Part II", in Occupational Psychology, 1956, Vol. 30, p. 57 - 68.

scale (z-scores) prior to their combination into a composite. The equation used was as follows.

$$\text{Composite} = \frac{\text{Position-Level Increase} + \text{Salary Increase}}{\text{Assessment Follow-up Time Interval}}$$

4. Procedures - Testing, Scoring, Data Gathering

The evaluation of nominees in the assessment centre program is conducted in groups, with twelve assesseees participating for a period of two and one-half days. This is followed immediately by two days during which the assessment information is reviewed and integrated into an overall assessment report. These in-residence programs are conducted in a training centre which provides for overnight accommodation. Each participant works on an individualized schedule centered on involvement in the various situational tests interspersed with the taking of diverse paper-and-pencil tests.

The paper-and-pencil instruments are administered in group situations in accordance with the standards and guidelines noted in the test manuals. This component of the assessment process is overseen by a trained test administrator. The Personality Research Form is taken by participants individually as part of a package completed just prior to attendance at the assessment centre. All psychometric paper-and-pencil tests are scored and rechecked by experienced trained scorers. Soon after each assessment program, the raw scores obtained on the various paper-and-pencil instruments are coded, keypunched and stored on magnetic tape for use in computer

analysis. The data are again double-checked at this stage.

With respect to the in-basket situational test, the scoring is conducted by graduate students in psychology in accordance with the extensive instructions delineated in the manual.¹¹⁸ The training of scorers in the in-basket technique is achieved by means of a comprehensive three-day training program. As with the paper-and-pencil tests, the dimensional scores from the in-basket are coded and stored for eventual computer analysis.

The assessors used to evaluate the eight performance dimensions in the other four situational tests are obtained from the executive levels of the federal government, supplemented at times with external assessors from private industry, provincial governments, and universities. Prior to their involvement as assessors, each executive participates in an intensive one-week training program. Operationally, a group of three assessors are assigned responsibility for assessing a group of six participants. For each simulation, one assessor assumes responsibility for observing and recording all relevant behavioural evidence, the result of which is a descriptive report of behaviour as it relates to each of the eight dimensions. Across the four simulations, an assessor will have observed a participant at least once, and frequently twice. Immediately following the two and one-half day assessment period, the three assessors come together

¹¹⁸ Senior Executive In-Basket Test: Scoring Manual, Personnel Psychology Centre, Public Service Commission of Canada, 1973.

for a day to integrate their observations of each participant.

During integration, all behavioural evidence related to a specific dimension is read and independently rated by each assessor. A seven point rating scale is used with a rating of four reflecting performance which meets the minimum level required for a senior executive. The assessors then discuss the assigned ratings, and this can lead to further attempts to rate performance based on additional information generated during discussion. Although an effort at consensus is encouraged, it is not deemed essential. This same process is used for each of the eight dimensions evaluated through these situational tests. The situational measures used in this study are the final average ratings assigned by the assessors to each of the eight dimensions. As with the other test measures, these dimensional scores are eventually coded and keypunched, and then stored on magnetic tape for analysis.

An important procedural point to note is the fact that all of the situational test measures are arrived at without there being any awareness of the performance of participants on the paper-and-pencil tests. Only after the assessors have determined the eight dimensional ratings are they then given information regarding a participant's performance on some paper-and-pencil tests, along with a descriptive summary and selected scores from the in-basket simulation. It is a combination of all of these sources of information which serves as the content of the overall assessment report.

It is also of importance to note that both the assessors and the in-basket test scorers have no access to any biographic or career-related background information which could potentially influence the observation and evaluation of performance in the situational tests. Moreover, as has been noted, no interview is conducted with participants by assessors as part of the assessment centre process. In fact, throughout the assessment centre program, various measures are taken to minimize and discourage even informal contact between the participants and the assessing staff.

With respect to the criterion measures employed for this study, the information was obtained from official records. These data were gathered over a three month period extending from December, 1976 to February, 1977. The time interval separating the collection of predictor and criterion measures ranged from approximately a year and one-half to four years. Once again, all such data were double-checked and transformed for computer analysis.

5. Statistical Analysis

In leading up to the testing of the Null Hypotheses generated for this study, some preliminary analyses were conducted, largely centered on descriptive statistics of the various predictor and criterion measures. As part of these analyses, estimates of skewness and kurtosis were obtained.¹¹⁹ The statistical significance of

¹¹⁹ Norman H. Nie, C. Hadlai Hall, Jean G. Jenkins, Karin Steinbrenner, and Dale H. Bent, Statistical Package for the Social Sciences, Second Edition, McGraw-Hill, Toronto, 1975, p. 184 - 185.

these measures were determined by means of tables presented by Croxton and Cowden.¹²⁰

The testing of all hypotheses required the use of a multiple regression procedure as a means of evaluating the contribution made by multiple predictors. For this purpose, a "simple" multiple regression model¹²¹ was employed. Such a procedure allows one to determine the amount of variance accounted for by a complete set of predictors. To test the significance of all multiple R's, Guilford and Fruchter's table of significant correlation coefficients was used.¹²²

In order to evaluate the first Null Hypothesis, that there is no difference in the variance accounted for by a paper-and-pencil set of predictors and a situational set of predictors, separate multiple R's were calculated for these two predictor sets to each criterion measure. Based on the assumption that the multiple R has the same predictive meaning as an ordinary bivariate correlation coefficient¹²³ the first Null Hypothesis was tested by means of a t test developed by

120 Frederick E. Croxton and Dudley J. Cowden, Applied General Statistics, Second Edition, Prentice-Hall Inc., Englewood Cliffs, 1955, p. 764 - 765.

121 Nie et al., Op.Cit., 1975, p. 320 - 367.

122 J.P. Guilford and Benjamin Fruchter, Fundamental Statistics in Psychology and Education, Fifth Edition, McGraw-Hill, Toronto, 1973, p. 515 - 516.

123 Quinn McNemar, Psychological Statistics, Fourth Edition, John Wiley and Sons, Toronto, 1969 p. 194.

Hotelling. This test compares the coefficients of correlation derived from correlated samples.¹²⁴

$$t_{dr} = (r_{12} - r_{13}) \sqrt{\frac{(N-3)(1+r_{23})}{2(1-r_{23}^2 - r_{12}^2 - r_{13}^2 + 2r_{23}r_{12}r_{13})}}$$

where r_{12} = correlation between a predictor set and criterion for one method of prediction;

r_{13} = correlation between a predictor set and criterion for a second method of prediction;

r_{23} = intercorrelation of the two predictor sets.

This formula necessitates an estimate not only of the correlation of the paper-and-pencil battery and of situational test battery to the criterion, but also of the paper-and-pencil predictors to the situational measures. This latter relationship was determined by obtaining the mean predictor score on each subject for both the paper-and-pencil and the situational measures, and correlating these mean scores. Since directionality is unknown, a two-tail test was called for in testing the hypothesis.

With respect to the second Null Hypothesis, that no incremental validity ensues from the addition of a set of paper-and-pencil measures to a set of situational test measures, the F formula taken from Guilford and Fruchter was used.¹²⁵

124 Guilford and Fruchter, Op. Cit., 1973, p. 167.

125 Ibid., p. 368.

$$F = \frac{(R_1^2 - R_2^2)(N - m_1 - 1)}{(1 - R_1^2)(m_1 - m_2)}$$

where R_1 = multiple R with the larger number of independent variables;

R_2 = multiple R with one (or more) variables omitted;

m_1 = larger number of independent variables;

m_2 = smaller number of independent variables.

This approach follows the suggestion of Marks, Christal and Bottenberg¹²⁶ who note that the extent to which a predictor(s) will increase the predictive efficiency of a set of predictors can be established by computing the multiple R both with and without the predictor(s) in question. Thus, for this second hypothesis multiple R's were calculated for the situational set of predictors and for the paper-and-pencil and situational sets combined. In the formula, the latter multiple R will represent the R with the larger number of variables, the former will represent the R containing the smaller number of variables.

The same procedure was used to test the third Null Hypothesis stating that no incremental validity ensues from the addition of a situational set of predictors to a set of paper-and-pencil measures. In this case the multiple R for paper-and-pencil predictors was compared with the multiple R established by combining both situational and paper-and-pencil measures.

126 M.R. Marks, R.E. Christal, and B.A. Bottenberg, "Simple Formula Aids for Understanding the Joint Action of Two Predictors", Journal of Applied Psychology, 1961, 45, p. 287.

CHAPTER 3

PRESENTATION AND DISCUSSION OF RESULTS

In this chapter, the results of the analyses of the data are presented. The chapter begins with a general description of the predictor and criterion variables employed in the study. In addition to the presentation of means, standard deviations and intercorrelation matrices, the scale qualities of skewness and kurtosis are noted for each of the variables. This is then followed by a presentation of the results relevant to the Null Hypotheses noted in the first chapter.

1. Descriptive Statistics and Intercorrelation Matrices for
the Paper-and-Pencil Predictors,
the Situational Predictors, and the Criterion Measures

- a. Paper-and-Pencil Test Measures:

Table I lists the means and standard deviations for the nine paper-and-pencil predictors employed in this study, along with measures of the distributional scale qualities of kurtosis and skewness. The intercorrelation of these measures is given in Table II. The order of inclusion of the predictors in Table II is the same as that noted in Table I. Inspection of the matrix table indicates that the intercorrelation among these paper-and-pencil predictors is in the low to moderate range. The median intercorrelation is .08 with a range extending from .01 to .53.

Table 1

Means, Standard Deviations and Measures of Kurtosis and Skewness
for the Nine Paper-and-Pencil Predictors
(N=297)

Variables	Mean	Standard Deviation	Kurtosis	Skewness
General Intelligence	66.92	12.01	0.12	-0.31
Administrative Judgement	27.15	4.65	0.12	-0.01
Achievement	15.91	2.40	-0.13	-0.41
Endurance	14.19	3.04	0.25	-0.39
Dominance	13.83	3.13	0.30	-0.45
Abasement	5.71	2.74	-0.22	0.10
Autonomy	8.41	2.77	-0.27	0.01
Succourance	7.33	2.98	-0.09	0.25
Affiliation	14.26	3.09	-0.17	-0.40

Significance at $p < .02$ requires a skewness score $\geq .11$

Significance at $p < .01$ requires a kurtosis score $\geq +.79$ or $-.54$

Table II
Intercorrelation Matrix of the Nine Paper-and-Pencil Predictors
(N=297)

Variables *	1	2	3	4	5	6	7	8	9
1	1.00								
2	0.42	1.00							
3	-0.04	0.02	1.00						
4	-0.06	-0.06	0.53	1.00					
5	-0.05	-0.07	0.34	0.32	1.00				
6	0.03	-0.03	0.12	0.10	-0.19	1.00			
7	0.01	-0.01	0.10	0.08	-0.11	0.01	1.00		
8	0.05	-0.04	-0.14	-0.21	-0.14	0.09	-0.48	1.00	
9	-0.04	-0.05	0.01	0.06	0.17	0.14	-0.25	0.25	1.00

* Variable 1: General Intelligence Test

Variable 2: Administrative Judgement Test

Variables 3 to 9 inclusive: PRF Dimensions

Significance at $p < .05$ requires $r \geq .113$

Significance at $p < .01$ requires $r \geq .148$

In terms of scale qualities, the measure of kurtosis will take on a positive value if the distribution of scores on a predictor is relatively peaked, a negative value if the distribution is flat in nature. Statistically significant evidence of kurtosis was not apparent for the paper-and-pencil predictors under consideration. However, several statistically significant relationships were in evidence for the measure of skewness. The skewness measure takes on a positive value when the distribution of cases on a variable tends to be clustered toward the left of the midpoint of the scale; a negative value indicates clustering of scores to the right or high end of a scale. Both for the measure of general intelligence and for various personality dimensions, the existence of a skewed distribution of scores was apparent. Five measures (intelligence, achievement, endurance, dominance and affiliation) showed a clustering of scores towards the high end of the scales, with one measure (succourance) showing a clustering towards the low end.

b. Situational Test Predictors:

In Table III a presentation is made of the descriptive statistics applicable to the fourteen situational measures of performance employed in this present study. The intercorrelation matrix of these predictors is noted in Table IV. In contrast with the low to moderate intercorrelations observed for the paper-and-pencil measures, the intercorrelation of the various situational predictors tends to be in the moderate to high range. As well, a definite pattern can be observed in Table IV in that the

Table III
Means, Standard Deviations, and Measures of
Kurtosis and Skewness for the Fourteen Situational Predictors
(N=297)

Variables	Means	Standard Deviation	Kurtosis	Skewness
Stress Tolerance	4.07	1.22	-0.65	-0.01
Motivation	4.34	1.31	-0.30	-0.13
Independence	3.88	1.33	-0.64	-0.01
Leadership	3.74	1.37	-0.74	0.01
Interpersonal Relations	4.12	1.26	-0.44	-.04
Analysis & Synthesis	3.93	1.35	-0.58	-0.01
Oral Communication	4.16	1.20	-0.51	-0.01
Quality of Judgement	3.59	1.22	-0.21	0.05
Initiating Activities	29.33	9.07	-0.03	0.15
Decision-Making	59.59	4.90	-0.39	-0.01
Problem Comprehension	59.96	4.80	-0.32	0.08
Delegation	51.29	1.54	0.60	0.10
Planning & Organization	24.15	10.85	-0.45	0.12
Productivity	78.03	16.12	-0.20	-0.06

Significance at $p < .02$ requires a skewness score $\geq .11$

Significance at $p < .01$ requires a kurtosis score $\geq +.79$ or $-.54$

Table IV
Intercorrelation Matrix of the Fourteen Situational Predictors
(N=297)

Variables *	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1.	1.00													
2.	0.58	1.00												
3.	0.61	0.67	1.00											
4.	0.69	0.69	0.68	1.00										
5.	0.55	0.54	0.39	0.62	1.00									
6.	0.58	0.69	0.68	0.67	0.55	1.00								
7.	0.72	0.60	0.54	0.67	0.59	0.60	1.00							
8.	0.60	0.65	0.61	0.66	0.61	0.79	0.62	1.00						
9.	0.25	0.30	0.33	0.38	0.28	0.31	0.26	0.34	1.00					
10.	0.19	0.18	0.21	0.24	0.12	0.19	0.12	0.20	0.43	1.00				
11.	0.10	0.17	0.22	0.23	0.18	0.30	0.13	0.24	0.48	0.39	1.00			
12.	0.10	0.16	0.22	0.20	0.07	0.14	0.09	0.17	0.23	0.55	0.20	1.00		
13.	0.15	0.27	0.20	0.27	0.25	0.27	0.14	0.27	0.44	0.30	0.36	0.18	1.00	
14.	0.24	0.30	0.31	0.32	0.28	0.30	0.24	0.34	0.81	0.42	0.54	0.25	0.45	1.00

* Variables 1 to 8 inclusive: Assessor Judgements
Variables 9 to 14 inclusive: In-Basket Scores

Significance at $p < .05$ requires $r \geq .113$

Significance at $p < .01$ requires $r \geq .148$

intercorrelation of the assessors' judgements tends to be higher than that noted for the predictors obtained through the in-basket simulation. Overall the range of correlations extends from .07 to .81 with a median intercorrelation of .30.

With respect to the distributional quality of kurtosis, the negative scores noted in Table III suggest a pattern towards a somewhat flat distribution of scores. Statistical significant evidence of kurtosis was obtained for the four dimensions of stress tolerance, independence, leadership, and analysis and synthesis. Somewhat unlike the paper-and-pencil predictors, there was an absence of a skewed distribution for most of the situational predictors. Only three of the fourteen measures showed statistical significant evidence of a clustering of scores away from the midpoint of the scale.

For the purpose of illustrating the general relationship between the situational and the paper-and-pencil predictors, Table V details the correlation coefficients obtained when the fourteen situational predictors were related to the nine paper-and-pencil measures. The correlations between the predictors of these two data sets tend, for the most part, to be quite low, suggesting little overlap between the two sets of predictors. The median correlation is .05 with a range extending from .01 to .41

c. Criterion Measures:

Table VI offers the descriptive and intercorrelation data relevant to the three criterion measures. As was noted in the second chapter, the position-level and salary criteria are moderately

Table V
Correlation of the Nine Paper-and-Pencil Predictors
with the Fourteen Situational Predictors
(N=297)

Variables*	1	2	3	4	5	6	7	8	9	10	11	12	13	14
a.	0.30	0.30	0.27	0.31	0.31	0.41	0.39	0.36	0.28	0.19	0.28	0.15	0.16	0.32
b.	0.21	0.10	0.18	0.15	0.11	0.21	0.25	0.18	0.23	0.12	0.26	0.14	-0.24	0.25
c.	-0.15	-0.09	-0.06	-0.14	-0.14	-0.08	-0.12	0.10	0.05	-0.09	-0.01	0.01	0.02	0.07
d.	-0.03	-0.05	-0.06	-0.04	-0.10	-0.06	-0.10	-0.08	0.02	-0.08	0.03	-0.06	0.05	0.01
e.	-0.04	-0.02	-0.07	-0.05	-0.02	-0.11	-0.09	-0.11	-0.05	-0.03	-0.02	-0.01	0.11	0.07
f.	0.04	-0.02	0.07	0.05	0.08	0.05	-0.03	0.03	0.05	-0.05	0.04	-0.06	0.04	0.02
g.	0.01	-0.06	0.01	-0.02	-0.05	0.02	-0.05	-0.10	-0.04	0.04	-0.01	0.10	0.05	-0.01
h.	0.01	0.03	0.03	0.01	0.02	0.01	-0.01	0.06	0.01	0.03	0.02	-0.03	0.04	-0.01
i.	0.05	0.04	-0.01	0.04	0.16	-0.03	0.04	0.04	0.03	-0.03	0.02	-0.08	0.03	0.05

* Variables 1 to 8 inclusive: Assessor Judgements

Variables 9 to 14 inclusive: In-Basket Scores

Variable a. General Intelligence Test

Variable b. Administrative Judgement Test

Variable c. to i. inclusive: PRF Dimensions

Significance at $p < .05$ requires $r \geq .113$

Significance at $p < .01$ requires $r \geq .148$

Table VI

Means, Standard Deviations, Intercorrelation Matrix, and
Measures of Kurtosis and Skewness for the three Criterion Measures
(N=297)

Variables*	1	2	3
1	1.00		
2	0.41	1.00	
3	0.79	0.81	1.00
Mean	0.40	6.78	0.04
Standard Deviation	0.43	7.72	0.69
Kurtosis	1.15	1.57	0.43
Skewness	0.76	-0.07	-0.01

* Variable 1: Rate of Position-Level Increase

Variable 2: Rate of Salary Increase

Variable 3: Composite of Variable 1 and 2

Significance at $p < .02$ requires a skewness score $\geq .11$

Significance at $p < .01$ requires a kurtosis score $\geq +.79$ or $-.54$

intercorrelated (.41) suggesting a degree of shared variance while at the same time possessing a fair proportion of unique variance. It is for this reason that a composite measure, combining both position-level and salary data, was incorporated into this study. The data in Table VI indicate that the rate of position-level increase and the rate of salary increase correlate with the composite measure at .79 and .81, respectively.

The mean of .40 noted for the criterion of position-level indicates that, on average, the Ss in this study progressed at a rate of one increase in level approximately every two and one-half years. The statistically significant measures of skewness and kurtosis further demonstrate that the distribution of scores on this criterion is relatively peaked toward the lower end of the scale. A major reason for this is that approximately one-third of the sample had made no upward progression in position-level at the time of follow-up, one and one-half to four years after assessment. Also, it is worthy to note that only five percent (N=15) of the sample had achieved the executive level of management at the time of follow-up. Thus, over the time period considered in this study, the upward progression of Ss had been somewhat restricted to the general middle and upper middle levels of management, levels which eventually lead to the executive target level.

With respect to the salary criterion, the mean score noted in Table VI shows that, on average, candidates increased their salary by six to seven hundred dollars per year beyond that which could be

attributed to general collective bargaining agreements. The large standard deviation points out the wide variability of scores; and the significant kurtosis measure suggests a distribution of scores which is relatively peaked. Appendix V presents frequency distribution tables for all of the predictor and criterion measures.

2. Statistical Testing of the Null Hypotheses

The statistical testing of the Null Hypotheses required the calculation of the simple multiple regression equations obtained for the paper-and-pencil predictors, the situational test measures, and for the paper-and-pencil and situational predictors combined. Table VII shows the simple Pearson r 's and multiple R 's which resulted when the nine paper-and-pencil predictors were related to the three criterion measures employed in this study. Similar information regarding the fourteen situational measures is presented in Table VIII. The order of inclusion of the variables into the regression equations is the same as listed in the respective tables. Appendix VI details the increment achieved with the addition of each predictor into the equation, along with the weights derived for each predictor measure.

When the paper-and-pencil and situational test measures were combined, a multiple R of .399 was obtained for the criterion of position-level. For the criterion of salary increase, the integration of both sets of predictors resulted in a multiple R of .291. The combining of the salary and position-level data into a composite led to a multiple R of .372. For these combined predictors, significance

Table VII

Pearson r 's and Multiple R 's Relating the Nine
Paper-and-Pencil Predictors to the
Three Criterion Measures
($N=297$)

Variables	Rate of Position- Level Increase	Rate of Salary Increase	Composite
General Intelligence	0.05	-0.03	0.01
Administrative Judgement	0.08	0.04	0.06
Achievement	0.01	-0.02	-0.01
Endurance	0.04	-0.06	-0.02
Dominance	0.03	0.08	0.05
Abasement	-0.09	-0.12	-0.12
Autonomy	-0.05	-0.06	-0.03
Succourance	-0.12	-0.06	-0.12
Affiliation	0.12	0.05	0.06
Multiple R	0.246	0.196	0.216

Significance at $p < .05$ requires $r \geq .113$

Significance at $p < .01$ requires $r \geq .148$

Significance at $p < .05$ requires $R \geq .223$

Significance at $p < .01$ requires $R \geq .253$

Table VIII

Pearson r's and Multiple R's Relating the Fourteen
Situational Predictors to the Three Criterion Measures
(N=297)

Variables	Rate of Position-Level Increase	Rate of Salary Increase	Composite
Stress Tolerance	0.04	-0.05	-0.02
Motivation	0.16	-0.03	0.09
Independence	0.14	0.04	0.09
Leadership	0.13	0.03	0.11
Interpersonal Relations	0.07	0.05	0.07
Analysis & Synthesis	0.17	0.05	0.16
Oral Communication	0.06	-0.02	0.01
Quality of Judgement	0.12	0.05	0.09
Initiating Activities	0.14	0.07	0.09
Decision Making	0.09	0.08	0.10
Problem Comprehension	-0.01	0.03	0.02
Delegation	0.07	0.07	0.09
Planning & Organization	0.10	0.06	0.10
Productivity	0.16	0.07	0.11
Multiple R	0.300	0.199	0.282

Significance at $p < .05$ requires $r \geq .113$

Significance at $p < .01$ requires $r \geq .148$

Significance at $p < .05$ requires $R \geq .258$

Significance at $p < .01$ requires $R \geq .287$

at the .05 level required a multiple R greater than or equal to .332; at the .01 level, a coefficient greater than or equal to .359 was needed.

The relevant data required for testing the first Null Hypothesis, that there is no difference in the amount of variance accounted for by a set of paper-and-pencil predictors and a set of situational test measures, is presented in Table IX. The testing of this hypothesis entailed the comparison of the multiple regression coefficients obtained in employing the paper-and-pencil predictors with the coefficients obtained in using the situational test measures. This comparison was made for each of the three criterion measures related to career progression. The statistical difference between these multiple regression coefficients was evaluated by means of Hotelling's t test which allows for the comparison of coefficients of correlations derived from correlated samples.¹²⁷ As has been noted, these analyses are based on the assumption that a multiple R has the same predictive meaning as an ordinary bivariate correlation coefficient.

Inspection of Table IX shows that across the three criterion measures, the multiple R's obtained for the paper-and-pencil predictors extends from .196 to .246. For the situational test predictors, the range is from .199 to .30. In terms of R^2 , the percentage of variance accounted for by the nine paper-and-pencil measures extends from 3.8% to 6.1%. The corresponding percentages for the fourteen

127 Guilford and Fruchter, Op. Cit., 1973, p. 167.

Table IX

Multiple R's and Statistical Tests of
Differences Obtained in Relating Paper-
and-Pencil Predictors and Situational
Predictors to Three Criterion Measures
(N=297)

Criterion Measure	Multiple R's Obtained for Paper-and-Pencil Predictors	Multiple R's Obtained for Situational Predictors	t_{dr}
Rate of Position-Level Increase	0.246	0.300	0.87
Rate of Salary Increase	0.196	0.199	0.05
Composite	0.216	0.282	1.06

Significance at $p < .05$ requires a $t \geq 1.97$

Significance at $p < .01$ requires a $t \geq 2.59$

predictors evaluated through situational tests is 4% to 9%.

The results of the t test conducted to compare the paper-and-pencil multiple R's with the corresponding situational multiple R's indicated a lack of statistical significance. This was in evidence for each of the three criterion measures. Thus, these data results indicate that the first Null Hypothesis can be accepted. It would appear that for the group of Ss considered in this study, there is no difference in the amount of variance accounted for by a set of situational test measures and a set of paper-and-pencil predictors.

Even though there is no statistical difference between the multiple R's obtained in using paper-and-pencil versus situational predictors, it may still occur that one set of predictors will add incrementally to the other in terms of predicting the career progression of participants to the assessment centre. As has been noted in the experimental design chapter one way of evaluating the incremental contribution made by a predictor or set of predictors is to compare the multiple R's obtained with and without the predictor variables in question. The statistical significance of this comparison is evaluated by the F test formula noted in Guilford and Fruchter.¹²⁸

The second Null Hypothesis was evaluated in this manner. This hypothesis states that there is no significant difference in the incremental validity added by a set of paper-and-pencil predictors

128 Ibid., p. 368.

when the validity associated with a set of situational predictors is taken into account. The multiple R's obtained both with and without the addition of the paper-and-pencil predictors are noted in Table X. The data presented in this table indicate that in using the criterion measure of position-level, a statistically significant F was obtained. The set of paper-and-pencil measures added incrementally to the variance accounted for by the situational measures in predicting the rate of position-level increase achieved by assessment centre participants one and one-half to four years following assessment. The magnitude of the multiple R increased from .30 to .399. In terms of the percentage of variance accounted for, this represents an increase from 9% to 15.9%.

When the salary criterion is isolated for consideration, the addition of the nine paper-and-pencil predictors served to increase the multiple R from .199 to .291. These coefficients reflect a percentage increase from 4% to 8.5%. Although this trend is somewhat similar to that noted for the position-level criterion, the F formula to test the difference between these two coefficients failed to attain statistical significance.

The results achieved for the composite measure of career progression parallels the position-level data, and a statistically significant F is present. When both the position-level and salary data are combined into a composite, the addition of a set of paper-and-pencil predictors results in a multiple R which increases from .282 to .372, from 8% to 13.8%.

Table X

Multiple R's and F Test of Differences
Obtained With and Without the Addition
of Paper-and-Pencil Predictors
(N=297)

Criterion Measure	Multiple R's Obtained Without Paper-and-Pencil Predictors	Multiple R's Obtained With Paper-and-Pencil Predictors Added	F
Rate of Position-Level Increase	0.30	0.399	2.51
Rate of Salary Increase	0.199	0.291	1.50
Composite	0.282	0.372	2.08

Significance at $p < .05$ requires $F \geq 1.91$

Significance at $p < .01$ requires $F \geq 2.48$

Taken as a whole, the results of these analyses indicate that the second Null Hypothesis is partially rejected. It would appear that for this group of assessment centre participants, a set of paper-and-pencil variables does contribute incrementally to the variance obtained with a set of situational measures of performance; however, the existence of this relationship is dependent to some extent on the criterion measure employed. In this study, the relationship was obtained using as a criterion measure the rate of position-level increase, as well as a composite measure combining both position-level and salary data. However, statistical significance was not attained in using the rate of salary increase as a measure of the progress achieved by assessment centre participants.

The relevant data involved in testing the third Null Hypothesis, that there is no difference in the incremental validity added by a set of situational test measures when the validity associated with a set of paper-and-pencil measures is taken into account, is presented in Table XI. As with the preceding hypothesis, the statistical operation to test this hypothesis involved a comparison of the multiple R's obtained when the situational predictors were added to the multiple regression equation and when they were not present in the equation.

In general, inspection of Table XI suggests a pattern somewhat similar to that which was noted in the testing of the preceding hypothesis. With respect to the criterion measure of rate of position-level increase, the data results indicate that the addition of

Table XI

Multiple R's and F Test of Differences
Obtained With and Without the Addition
of Situational Predictors
(N=297)

Criterion Measure	Multiple R's Obtained Without Situational Predictors	Multiple R's Obtained With Situational Predictors Added	F
Rate of Position-Level Increase	0.246	0.399	2.30
Rate of Salary Increase	0.196	0.291	0.98
Composite	0.216	0.372	2.08

Significance at $p < .05$ requires an $F \geq 1.73$

Significance at $p < .01$ requires an $F \geq 2.15$

performance measures obtained through situational tests adds in an incremental fashion to the variance accounted for by a set of paper-and-pencil measures. The addition of the fourteen situational predictors resulted in increasing the multiple R from .246 to .399. In terms of R^2 , the percentage of variance accounted for in the predictor-criterion relationship increased from 6.1% to 15.9%, an increase of approximately ten percent. The F statistic calculated to test this difference in multiple R's was found to be statistically significant.

The pattern of results was the same when the composite index was employed as the criterion measure. In this case, the significant F statistic supported the observation showing that the multiple R increased from .216 to .372. The incorporation of the situational measures served to increase the percentage of variance from 4.7% to 13.8%, an increase of nine percent. In using the rate of salary increase as the criterion measure, the multiple R increased from .196 to .291; however, as with the preceding hypothesis, this increase in the magnitude of the coefficient was not found to be statistically significant.

Overall, these results indicate that the third Null Hypothesis, that a set of situational predictors does not add incrementally to the validity accounted for by a set of paper-and-pencil predictors, can be rejected in part. Without the addition of the situational measures of performance, the percentage of variance accounted for extended from 4.7% to 6.1%. The addition of these situational

variables resulted in percentages ranging from 8.5% to 15.9%. The Null Hypothesis is accepted for the criterion measure of rate of salary change. However, it is rejected for the criterion measure of rate of position-level increase, as well as for the composite index incorporating both salary and position-level measures of the general career progression achieved by assessment centre participants.

Within the context of the assessment centre program conducted for the Canadian Federal Public Service, this summary of research findings supports the following general conclusions.

1. No statistically significant difference was found in terms of the amount of predictive variance accounted for by a set of paper-and-pencil predictors and by a set of situational test predictors. In using a simple multiple regression model as an evaluative tool, this lack of a statistical difference between the two sets of predictors was found for each of the three criterion measures of general career progression within an organization: rate of position-level increase, rate of salary increase, and a composite incorporating both position-level and salary data. Across the three criterion measures, the multiple R's obtained in using the paper-and-pencil predictors extended from .196 to .246; in employing the situational predictors, the multiple R's ranged from .199 to .30.

2. Evidence was gathered to support the observation that a set of paper-and-pencil predictors contributed incrementally to the overall prediction of organizational progression. Once the variance attributed to the set of situational test variables was taken into account, it was found that the paper-and-pencil predictors added to this in an incremental fashion. Although this pattern was noted for all three criterion measures, statistical significance was achieved only for the criterion measure of position-level and for the composite index. In the case of the position-level criterion, the addition of paper-and-pencil variables resulted in a multiple R

of .399, in contrast with the multiple R of .30 obtained with the situational predictors alone. For the composite index, the multiple R increased from .282 to .373.

3. A similar pattern was evident when the incremental validity associated with a set of situational test measures was explored. The results of these analyses showed that for the criterion measure of position-level and for the composite index, the situational test predictors contributed incrementally to the variance accounted for by a set of paper-and-pencil predictors. For these two criterion measures, the addition of situational test scores served to increase the multiple R's from .246 to .399, and from .216 to .373, respectively. This relationship failed to achieve statistical significance when the rate of salary increase was used as a criterion measure.

Overall, the results of this investigation indicate that the amount of predictive variance accounted for by a set of paper-and-pencil predictors and by a set of situational predictors tends to be similar. At the same time, both methods of assessment can be found to add incrementally to each other in an overall prediction of career progression within an organization, although this relationship is to some extent dependent upon the criterion measure employed. The results do tend to support the various comments made regarding the premature rejection of psychometric paper-and-pencil tests in favour of situational tests. These findings of course apply to the type of population undergoing assessment within the context of the assessment centre program conducted for the Canadian Federal Public Service. In this study it was possible to follow-up 92% of the assessed population.

Nonetheless, caution should be exercised in the generalization of these results. On the one hand, the population involved in this research was heterogeneous with regard to various biographic and career-related variables including age, occupational classification, length of service, organizational level, and so forth. On the other hand, some prescreening of participants did occur in that the various government departments were encouraged to send only participants showing the greatest potential for performance at executive levels of management within the organization.

Again, the findings reported apply strictly to the specific predictor and criterion variables and multiple regression strategies utilized in this research. Some interesting extensions to this investigation can be noted. Continued follow-up of this and other samples over a longer time frame should prove advantageous. As has been noted, a sizeable proportion of the assessment centre participants had not made any upward progression in the organization over the one and one-half to four year follow-up time interval. This occurred despite the fact that all such participants were viewed within their respective departments to have high potential for effectiveness at executive levels of management. As well, only a small percentage of subjects were able to achieve the targeted executive levels during the assessment follow-up time interval. These factors suggest that a relatively short time-frame may have been involved in this investigation. It may occur that the results obtained in this study will become accentuated over time as the assessment

centre participants distribute themselves across a wider range of organizational levels and salary bands, including more subjects who would eventually achieve the executive levels of management.

Such long-term longitudinal research also holds the advantage of allowing for the investigation of the dynamic character of the predictor-criterion relationship. The writings of Ghiselli^{129, 130} have served as a prime impetus for research in this area. In part, it is his contention that predictor-criterion relationships may increase or decrease with time, or alternatively, may result in a kind of cyclical pattern. Consideration of the changing rather than static nature of managerial criterion has also been given greater emphasis within various validation models and strategies.^{131, 132}

Within the assessment centre area, a few recent investigations have begun to give more consideration to this factor. For example, a study by Mitchel¹³³ showed that the predictive validity of assessment centre predictors tended to increase over time. Additional research

129 Edwin E. Ghiselli, "Dimensional Problems of Criteria," in Journal of Applied Psychology, Vol. 40, 1956, p. 1 - 4.

130 Edwin E. Ghiselli and Mason Haire, "The Validation of Selection Tests in the Light of the Dynamic Character of Criteria", in Personnel Psychology, Vol. 13, 1960, p. 225 - 231.

131 Lawrence R. James, "Criterion Models and Construct Validity for Criteria", in Psychological Bulletin, Vol. 80, 1973, p. 75 - 83.

132 Paul F. Wernimont and John P. Campbell, "Signs, Samples, and Criteria", in Journal of Applied Psychology, Vol. 52, 1968, p. 372 - 376.

133 Mitchel , Op. Cit., 1975, p. 573 - 579.

by the present author ¹³⁴ has also demonstrated that, dependent upon the criterion considered, the existence of a predictor-criterion relationship may actually be time-locked or pinpointed in time. One extension of the research obtained in the present investigation would be to observe the possible dynamic aspect of criteria as they apply to both paper-and-pencil and situational methods of assessment. Only long-term longitudinal research of predictive validity can offer tentative responses in this important area.

A second extension to the present research would involve the use of a variety of criterion measures, in order to tap the many facets of what may be considered under the rubric of organizational or managerial "success". Reflecting the general criterion validity approach of other assessment centre researches, the present investigation emphasized the use of global and composite measures of organization outcomes. Extensions to this basic design are possible by establishing the individual and incremental predictive validity of situational and paper-and-pencil testing methods across a greater diversity of criterion measures. As guiding models in this area, some authors ^{135,136} have pointed out the need to explore job

134 Slivinski et al, Op. Cit., 1978, p. 6 - 11.

135 L.W. Slivinski and K.W. Grant, "Behavioural Validation Prediction Model", Paper presented to the International Personnel Association, San Francisco, 1971, p. 1 - 39.

136 Wernimont and Campbell, Op. Cit., 1968, p. 372 - 376.

performance criteria as a means of establishing the consistency which exists between measures of behaviour arrived at through predictor tests and the same or similar behaviours evaluated on the job. An acknowledged difficulty with these behavioural consistency models is that such evaluations may reflect more the measurement of reliability than the measurement of validity.¹³⁷ Recent research investigations by the present author,^{138,139} have begun to investigate the relationship between assessment centre predictors and similar job performance behaviours.

Consideration of other criterion indices are also possible. In the most recent critical review of the assessment centre approach, Klimoski and Struckland¹⁴⁰ have argued that the full validity and utility of various assessment centre methodologies can only emerge through consideration of other criterion measures in addition to those related to general career advancement. As one alternative, they suggest measures designed to reflect the "value" of individuals to the organization. Included here would be such criteria as an estimate of the replacement cost of an individual, measures of the

137 Ibid., p. 373.

138 Slivinski et al., Op. Cit., 1978, p. 6 - 11.

139 L.W. Slivinski and J.L. McCloskey, "Canadian Middle Management Assessment Centre", Paper presented at the Sixth International Assessment Centre Congress, Greenbrier, April, 1978.

140 Klimoski and Struckland, Op. Cit., 1977, p. 353 - 361.

management objectives achieved, and measures of the critical contribution made by an individual to the organization.

From a somewhat different perspective, James ¹⁴¹ points out that an additional source of criterion information could be arrived at through the quantification of "job behaviour" measures centered on what a manager "actually does" in performing a job. As has been noted, the subjects used in the present investigation were quite heterogeneous in terms of their occupational and managerial duties, responsibilities and functions both at the time of assessment and at the time of follow-up. The criterion analysis of job behaviours would be one methodological means through which the type and degree of management involved in a job could be established. The question arises as to what the relative predictive validities would be for paper-and-pencil and situational predictors when consideration is given to the managerial job behaviours, roles, and responsibilities of the subjects.

In fact, exploration of this type of area points out a third major extension to the present research, namely, the analysis of the possible differential predictive validities which might be established with more homogeneous subgroups of subjects. The test validation

141 James, Op. Cit., 1973, p. 75 - 83.

model presented by Hobert and Dunnette,¹⁴² among others,^{143,144} emphasizes that a number of intervening variables operate to influence any predictor-criterion relationship. The concept of job behaviours, that is, the type and degree of management involved in a job, would be one such intervening variable. Other variables would include the different groups or types of individuals involved, as well as a multitude of situational factors. In order to give consideration to the complex interactions existing between predictors and criteria, it becomes necessary to identify more homogeneous subgroups of individuals, job behaviours, and situations. It may occur that the pattern of relative and incremental predictive validities established in this study for a fairly heterogeneous group of managerial aspirants will alter and vary with the analysis of more homogeneous subgroups.

Still a fourth possible extension to the present investigation would be to explore the predictive validities associated with other assessment methods, particularly those methods which are so frequently used as part of an assessment centre approach. The focus

142 Robert Hobert and Marvin Dunnette, "Development of Moderator Variables to Enhance the Prediction of Managerial Effectiveness", in Journal of Applied Psychology, Vol. 51, 1967, p. 50 - 64.

143 Frank L. Schmidt and Leon B. Kaplan, "Composite vs. Multiple Criteria: A Review and Resolution of the Controversy", in Personnel Psychology, Vol. 24, 1971, p. 419 - 434.

144 Wernimont and Campbell, Op. Cit., 1968, p. 372 - 376.

of this study was to establish the relative and incremental validities of a set of paper-and-pencil predictors and a set of situational predictors. A natural extension would be to explore further the unique and incremental contribution of each of the various assessment tools which serve to form the general process known as the assessment centre. Certainly prime amongst these would be the role of the selection interview and of various biographic and career-related measures. Only in this way can the uniqueness of the assessment centre process be established.

Cognizance must be made of the fact that the operational use of paper-and-pencil tests and/or situational tests will be determined not only by various validity estimates, but also by such factors as cost, acceptability, and utility. Further continued research is needed to determine how each of these factors impinge upon the various assessment methodologies. Some authors,^{145,146} have noted that the incremental or superior validity so frequently attributed to the assessment centre approach (and implicitly to situational tests) must be tempered with the expensive cost involved in using such an approach.

In addition to costing, the general acceptability of an evaluative tool is of obvious importance and it is perhaps for this

145 Klimoski and Struckland, Op. Cit., 1977, p. 353 - 361.

146 Wilson and Tatge, Op. Cit., 1973, p. 172 - 179.

reason that there has been an almost "bandwagon" effect¹⁴⁷ in the use of a situational approach to managerial selection. Certainly much of the attitudinal reasearch to date strongly supports the high acceptability of situational tests by assessors, by candidates, and by organizational management.¹⁴⁸ It is of interest to note, however, that attitudinal research conducted by the present author has clearly demonstrated that when paper-and-pencil testing approaches are appropriately explained and discussed, their general acceptability can frequently be exceedingly high.¹⁴⁹ In this area, it must also be acknowledged that within some countries, particularly the United States, the general assessment centre method has received a degree of legal acceptability, in large measure due to the fact that the development of situational tests emphasizes the establishment of test instruments which are clearly job-related and job-relevant.¹⁵⁰

Finally, the ultimate use of psychometric paper-and-pencil and/or situational tests within any organizational or managerial

147 Robert B. Finkle, "Managerial Assessment Centres", in Marvin D. Dunnette (Ed.), Handbook of Industrial and Organizational Psychology, Chicago, Rand McNally, 1976, p. 861 - 888.

148 William E. Dodd, "Attitudes Toward Assessment Center Programs", in Joseph L. Moses and William C. Byham (Eds.), Applying the Assessment Center Method, Toronto, Pergamon Press, 1977, p. 161 - 184.

149 L.W. Slivinski, T.R. Yan, I.L. Richter, V.S. McDnald and R.P. Bourgeois, Reactions to an Assessment Centre, Technical Report CAP-AC 28, Personnel Psychology Centre, Public Service Commission, 1977, 1 - 77 p.

150 William C. Byham, "Application of the Assessment Center Method", in Joseph L. Moses and William C. Byham (Eds.), Applying the Assessment Center Method, Toronto, Pergamon Press, 1977, p. 31 - 43.

context will be influenced by the general utility of the information which is gathered. Certainly, one major criticism of the psychometric tradition has been that the data collected lacks "decisional utility".¹⁵¹ However, the utility of all test information can be looked at from many different perspectives. In addition to evaluating test instruments in terms of their usefulness in facilitating selection or promotion decisions, one may also speak of the utility within the context of training and development strategies, career planning systems, and individual counselling and feedback processes. In all of these areas, both psychometric paper-and-pencil instruments and situational tests, as well as other evaluative tools, may hold high potential.

An additional point with regards to this research concerns itself with cross-validation. As pointed out by Claudy¹⁵² there are three outcomes that a researcher is seeking when applying a multiple regression to a set of data. The outcome that is appropriate to this research is to determine the predictive effectiveness of the regression equation when it is applied to future new cases from the population. The extension of this thesis then, would be to cross-validate these obtained results with a second independent sample from the same population at a later time to obtain a predicted dependent variable value for each criterion observation in the second sample.

151 Bersoff, Op. Cit., 1973, p. 892 - 899.

152 John G. Claudy, "Multiple Regression and Validity Estimation in One Sample", in Applied Psychological Measurement, Vol. 2, 1978, 595-607.

BIBLIOGRAPHY

Administrative Judgement Test Manual, U.S. Civil Service Commission, Technical Series, No. 34, 1966, 1 - 15 p.

This manual details the development, research, and use of the Administrative Judgement Test which was used as one of the paper-and-pencil test measures in this study.

Bentz, V. Jan, "Validity of Sears Assessment Centre Procedures", Paper presented at American Psychological Association, Washington, 1971.

This presentation offers extensive statistical data on various predictor-criterion relationships within the assessment center context.

Bersoff, Donald N., "Silk Purses into Sow's Ears - The Decline of Psychological Testing and a Suggestion for its Redemption", in American Psychologist, Vol. 28, 1973, p. 892 - 899.

A critical appraisal of general psychological testing leading to the suggestion of a situational approach to testing as an alternative.

Bourgeois, R.P. and Slivinski, L.W., "The Inter-rater Reliability of the Consolidated Fund In-Basket", in Studies in Personnel Psychology, Vol. 6, 1974, p. 47 - 52.

This article presents reliability estimates for an in-basket test which served as a precursor to the one employed in this study.

Bray, Douglas W. and Campbell, Richard J., "Selection of Salesmen by Means of an Assessment Centre", in Journal of Applied Psychology, 1968, Vol. 52, p. 36 - 41.

A predictive validity study of the assessment centre method which compares in part, the validity of the overall assessment judgement with the validity associated with paper-and-pencil cognitive ability tests.

-----, Campbell, Richard J. and Grant, Donald L., Formative Years in Business: A Long-term AT&T Study of Managerial Lives, New York, Wiley, 1974, XVIII - 236 p.

A general text of the first major industrial-organizational use of the assessment centre method, offering a global longitudinal perspective of the success and progression of the assessment participants.

----- and Grant, Donald L., "The Assessment Centre in the Measurement of Potential for Business Management", in Psychological Monographs, Vol. 80, 1966. Whole No. 625, p. 1 - 27.

The widely-quoted major predictive validity study of the first use of an assessment centre within a managerial context.

Byham, William C., "Application of the Assessment Center Method", in Joseph L. Moses and William C. Byham, (Eds.), Applying the Assessment Center Method, Toronto, Pergamon Press, 1977, p. 31 - 43.

A general article of current assessment centre applications, including comments on the job relevance and job relatedness of the method within the context of various legal guidelines.

Campbell, John P., Dunnette, Marvin D., Lawler, Edward E. and Weick, Karl E., Managerial Behaviour, Performance, and Effectiveness, New York, McGraw-Hill, 1970, XIV - 546 p.

As part of a global examination of current practices used in identifying managerial talent, the authors review and comment on several major paper-and-pencil validation research projects.

Carleton, F.O., "Relationships Between Follow-up Evaluations and Information Developed in a Management Assessment Centre", Paper presented at the American Psychological Association, 1970.

This paper presents predictive validity coefficients obtained in relating various assessment centre components, including an interview and projective test, with follow-up criteria including supervisory ratings and a composite index of career progression.

Cohen, Barry, M., Moses, Joseph L. and Byham, William C., The Validity of Assessment Centres: A Literature Review, Monograph II, Development Dimensions Inc., Pittsburg, 1974, 1 - 31 p.

A comprehensive and detailed review of all available validation studies conducted on the assessment centre method up to and including the year 1972, pointing out the consistency with which the assessment centre judgement predicts career progression and job performance.

Dodd, William E., "Attitudes Toward Assessment Center Programs", in Joseph L. Moses and William C. Byham (Eds.), Applying the Assessment Center Method, Toronto, Pergamon Press, 1977, p. 161 - 184.

A summary of some current research tapping the mostly positive attitudes of participants and managers toward the assessment centre process.

Dunnette, Marvin D., "The Assessment of Managerial Talent", in Paul McReynolds (Ed.), Advances in Psychological Assessment, Palo Alto, Science and Behaviour Books, Vol. II, 1971, p. 79 - 108.

A general review and perspective of assessment centre research, emphasizing in part the predominant role of situational tests and the incremental validity resulting from the use of multiple predictors.

-----, "Personnel Management", in Annual Review of Psychology, Vol. 8, 1962, p. 285 - 314.

A general perspective on the status of personnel management research leading to the suggestion of a moratorium on the development of further tests.

Claudy, J.G., "Multiple Regression and Validity Estimation in One Sample", in Applied Psychological Measurement, Vol. 2, 1978, p. 595-607.

Finkle, Robert B., "Managerial Assessment Centers", in Marvin D. Dunnette (Ed.), Handbook of Industrial and Organizational Psychology, Chicago, Rand McNally, 1976, p. 861 - 888.

A general review of assessment centre programs, pointing out in part the almost bandwagon approach in its use, and the need for continued professional direction and guidance.

Finley, Robert M., "Evaluation of Behaviour Predictions from Projective Tests Given in a Management Assessment Centre", Paper presented at the American Psychological Association, 1970.

A research project oriented to establishing the validity of a projective test component of an assessment centre.

General Intelligence Test Manual (Second Edition), Personnel Psychology Centre, Public Service Commission of Canada, 1977, 1 - 40 p.

This manual details the background development and current research data available on the General Intelligence Test which was employed as a paper-and-pencil test measure for this investigation.

Ghiselli, Edwin E., "Dimensional Problems of Criteria", in Journal of Applied Psychology, Vol. 40, 1956, p. 1 - 4.

One of the initial articles to speak of the dynamic character of predictor-criterion relationships.

-----, The Measurement of Occupational Aptitude, Los Angeles, University of California Publications, 1955, 101 - 216 p.

A major attempt to bring together and summarize all validity studies related to the identification and selection of personnel in organizations.

-----, "The Validity of Aptitude Tests in Personnel Selection", in Personnel Psychology, Vol. 26, 1973, p. 461 - 477.

An article which updates previous writings related to the occupational validity of selection tests for personnel selection.

----- and Haire, Mason, "The Validation of Selection Tests in the Light of the Dynamic Character of Criteria", in Personnel Psychology, Vol. 13, 1960, p. 225 - 231.

A research investigation illustrating the changing pattern which exists over time in predictor-criterion relationships.

Glaser, Robert, Schwarz, Paul A. and Flanagan, John C., "The Contribution of Interview and Situational Performance Procedures to the Selection of Supervisory Personnel", in Journal of Applied Psychology, Vol. 42, 1958, p. 68 - 73.

One of the few articles showing the relative and incremental contribution made by situational and paper-and-pencil tests in the selection of supervisory personnel.

Goldfried, Marvin R. and Kent, Ronald N., "Traditional Versus Behavioural Personality Assessment: A Comparison of Methodological and Theoretical Assumptions", in Psychological Bulletin, Vol. 77, 1972, p. 409 - 420.

This article compares the traditional and behavioural or situational approaches to the assessment of personality, including an analysis of various methodological issues and underlying assumptions.

Grant, Donald L. and Bray, Douglas W., "Contributions of the Interview to Assessment of Management Potential", in Journal of Applied Psychology, Vol. 53, 1969, p. 23 - 34.

As part of a series of articles on the first organizational use of an assessment centre, this article attempts to isolate the contribution made by the interview component of the process.

-----, Katkovsky, Walter and Bray, Douglas W., "Contributions of Projective Techniques to Assessment of Management Potential", in Journal of Applied Psychology, Vol. 51, 1967, p. 226 - 232.

Another research investigation of the first organizational use of an assessment centre, emphasizing the role to be played by projective tests of personality.

Grimsley, Glen and Jarrett, Hilton F., "The Relation of Past Managerial Achievement to Test Measures Obtained in the Employment Situation: Methodology and Results", in Personnel Psychology, Vol. 26, 1973, p. 31 - 48.

Using a concurrent validation strategy, this research represents one of the most recent attempts to show the predictive value of a battery of paper-and-pencil test measures.

----- and Jarrett, Hilton, F., "The Relation of Past Managerial Achievement to Test Measures Obtained in the Employment Situation: Methodology and Results-II", in Personnel Psychology, Vol. 28, 1975, p. 215 - 231.

This article represents an extension of the above-noted research in which the generalizability and replicability of findings is established by determining the validity of multiple paper-and-pencil predictors for a second sample of management personnel.

Guilford, J.P. and Fruchter, Benjamin, Fundamental Statistics in Psychology and Education, Fifth Edition, Toronto, McGraw-Hill, 1973, X - 546 p.

A source book for statistical procedures.

Guion, Robert M. and Gottier, Richard F., "Validity of Personality Measures in Personnel Selection", in Personnel Psychology, Vol. 18, 1965, p. 135 - 164.

A review of the validity of paper-and-pencil tests of personality and interest leading to relatively pessimistic conclusions on the status of such psychometric tests.

Hedlund, Dalia E., "A Review of the MMPI in Industry", in Psychological Reports, Vol. 17, 1965, p. 875 - 889.

This review of research shows the poor predictive value of one particular personality test for the purpose of selecting personnel.

Hobert, Robert and Dunnette, Marvin, "Development of Moderator Variables to Enhance the Prediction of Managerial Effectiveness", in Journal of Applied Psychology, Vol. 51, 1967, p. 50 - 64.

This representative article points out the complexity of all predictor-criterion relationships, particularly how they are influenced by a number of intervening variables.

Howard, Ann, "On Assessment of Assessment Centers", in Academy of Management Journal, Vol. 17, 1974, p. 115 - 134.

A review of assessment centre research in which the author concludes that more research is necessary to determine the contribution of various assessment centre components, particularly the role of psychometric paper-and-pencil tests.

Huck, James R., "Assessment Centers: A Review of the External and Internal Validities", in Personnel Psychology, Vol. 26, 1973, p. 191 - 212.

A review of assessment centre validation studies, pointing out the critical need for more research on the contribution made by the various assessment procedures.

----- and Bray, Douglas W., "Management Assessment Centre Evaluations and Subsequent Job Performance of White and Black Females", in Personnel Psychology, Vol. 29, 1976, p. 13 - 30.

A predictive validity study demonstrating the relationship which exists between the overall assessment centre judgement and subsequent job performance.

James, Lawrence R., "Criterion Models and Construct Validity for Criteria", in Psychological Bulletin, Vol. 80, 1973, p. 75 - 83.

A representative validation model which speaks of the dynamic aspect of criteria, as well as of the need to explore job behaviours as criterion measures.

Jackson, Douglas N., Personality Research Form Manual, Goshen, Research Psychologists Press, 1967, 1 - 31 p.

The manual details the development and research conducted on the paper-and-pencil personality test used, in part, for the present research investigation.

Klimoski, Richard J. and Struckland, William J., "Assessment Centres: Valid or Merely Prescient", in Personnel Psychology, Vol. 30, 1977, p. 353 - 361.

The most recent critical review of the assessment centre method emphasizing the need for research designed to compare various assessment methods of selection, as well as the need to explore a variety of criterion measures.

Korman, Abraham K., "The Prediction of Managerial Performance: A Review", in Personnel Psychology, Vol. 21, 1968, p. 295 - 322.

A review and evaluation of predictive validity studies related to the selection of personnel within organizations, including various psychometric paper-and-pencil tests of cognitive ability, personality, and interests.

Laurent, Harry, "Early Identification of Managers", in Management Record, May, 1962, p. 33 - 38.

A description of the development and initial validation research conducted on an organizational program which utilized a variety of paper-and-pencil test measures to select and identify management personnel.

-----, "Research on the Identification of Management Potential", in John A. Myers (Ed.), Predicting Managerial Success, Ann Arbor, Braun and Brumfield, 1968.

Further validation research on the identification of management potential through the use of multiple paper-and-pencil test predictors, including some favourable results obtained when used with non-management personnel.

-----, "Cross-Cultural Cross-Validation of Empirically Validated Tests", in Journal of Applied Psychology, Vol. 54, 1970, p. 417 - 423.

Continued research of the validity of a paper-and-pencil test battery, demonstrating the validity estimates obtained across samples from different countries.

MacKinnon, Donald W., "Human Assessment: Perspective and Context for Current Practice", Invited address to American Psychological Association, 1975, 1 - 31 p.

An important perspective of the assessment centre process, arguing against the premature rejection of psychometric test instruments and leading to the suggestion that more research be conducted on the relative contribution of paper-and-pencil and situational approaches to assessment.

Marquardt, Llyod D., Follow-up Evaluation of the Second Look Approach to the Selection of Management Trainees, Chicago, Sears/Roebuck, 1976.

This article describes a follow-up evaluation of various aspects of an assessment centre, perhaps most noted for the use of a large number of criterion measures.

McClelland, David C., "Testing for Competence Rather than Intelligence", in American Psychologist, Vol. 28, 1973, p. 1 - 14.

The author critiques the traditional methods of testing intelligence and aptitudes, and this leads to the suggestion of a criterion-sampling approach.

McDonald, Vince and Ricard S., "Intra-class Correlations Among Assessors' Integration Ratings", Unpublished Document, 1977, 1 - 6 p.

This article offers reliability estimates for the assessor judgements which are utilized in this study as part of the situational test measures.

Mitchel, James O., "Assessment Centre Validity: A Longitudinal Study", in Journal of Applied Psychology, Vol. 60, 1975, p. 573 - 579.

A predictive validity study of an assessment centre, noted in part for its longitudinal approach in looking at the predictor-criterion relationship over a five year time span.

Moses, Joseph L., "Assessment Centre Performance and Management Progress", in Studies in Personnel Psychology, Vol. 4, 1972, p. 7 - 12.

As part of this validation study of the assessment centre process, the author compares the validity of cognitive test measures with that obtained using the overall assessment judgement of potential.

----- and Boehm, Virginia R., "Relationship of Assessment Center Performance to Management Progress of Women", in Journal of Applied Psychology, Vol. 60, 1975, p. 527 - 529.

This study demonstrated the relationship which exists between the overall assessment centre rating of potential and the career progression made by participants two to ten years later.

Nash, Allan N., "Vocational Interests of Effective Managers, A Review of the Literature", in Personnel Psychology, Vol. 18, 1965, p. 21 - 37.

This literature review concluded that some evidence was available to indicate that interest patterns could be related to effectiveness within a management context.

Nie, Norman H., Hall, Hadlai, Jenkins, Jean G., Steinbrenner, Karin and Bent, Dale H., Statistical Package for the Social Sciences, Second Edition, McGraw-Hill, Toronto, 1975, XXIV - 675 p.

A manual used in conjunction with SPSS computer program package.

Office of Strategic Services (OSS) Assessment Staff, Assessment of Men, New York, Rinehart, 1948.

A historical text detailing the development and operation of the first major North American use of the assessment centre method within a military setting.

Prien, Erich, "The In-Basket Test: A Review of Research and Opinion", Unpublished Document, 1977.

The most comprehensive and detailed review of the in-basket technique to date, including such aspects as design and development, applications, reliability estimates, validity studies, and correlates of performance.

Schmidt, Frank L. and Kaplan, Leon B., "Compcsite vs. Multiple Criteria: A Review and Resolution of the Controversy", in Personnel Psychology, Vol. 24, 1971, p. 419 - 434.

A representative general validation strategy emphasizing in particular the complexity of the problems associated with criterion measures.

Senior Executive In-Basket Test: Scoring Manual, Personnel Psychology Centre, Public Service Commission of Canada, 1973.

A manual used to score the in-basket situational test used in this research investigation.

Slivinski, L.W., "The Relation of the Supervisory Judgement Test to Career Progression and Job Performance", Unpublished Data, 1977.

These data illustrate the relationship between a paper-and-pencil test of supervisory judgement and eventual progression and performance in first-line supervisory positions.

----- and McCloskey, J.L., "Canadian Middle Management Assessment Centre", Paper presented at the Sixth International Assessment Centre Congress, Greenbrier, April, 1978.

This predictive validity research explores a variety of criterion measures of career progression and job performance, including an attempt to establish the relationship between various dimensions measured at the assessment centre and the evaluation of the same dimensions assessed on the job.

-----, Bourgeois, R.P., Pederson, L.P. and McCloskey, J.L., Development and Application of the Assessment Centre for the Identification of Senior Executive Potential, Public Service Commission of Canada, In Press, 1978.

This comprehensive report details all aspects of the development and operation of the assessment centre program which was employed to evaluate the participants used as subjects for this present investigation.

----- and Grant, K.W., "Behavioural Validation Prediction Model", Paper presented to the International Personnel Association, San Francisco, 1971, 1 - 39 p.

A representative validation strategy noted for its emphasis on linking assessment centre behaviours with identical or similar on-the-job behaviours.

-----, Grant, K.W., Bourgeois, R.P., Pederson, L.D. and McCloskey, J.L., "Longitudinal Follow-up of a First Level Management Assessment Centre", in Journal of Assessment Center Technology, Vol. 1, 1978, p. 6 - 11.

A predictive validity study relating assessment centre performance to global, composite, and multiple criterion measures obtained at three different points in time including the establishment of the link between specific assessment centre behavioural dimensions and similar job performance behaviours.

-----, Yan, T.R., Richter, I.L., McDonald, V.S. and Bourgeois, R.P., Reactions to an Assessment Centre, Technical Report CAP-AC 28, Personnel Psychology Centre, Public Service Commission, 1977, 1 - 77 p.

A summary article bringing together several reports, all of which demonstrate the positive reactions of both assessors and participants to the assessment centre conducted for the Federal Public Service of Canada.

Sluckin, W., "Combining Criteria of Occupational Success: Part I", in Occupational Psychology, 1956, Vol. 30, p. 20 - 26.

A general article on the problems and methods associated with the use of criteria including some general guidelines for the combining of criterion measures.

-----, "Combining Criteria of Occupational Success: Part II", in Occupational Psychology, 1956, Vol. 30, p. 57 - 68.

An extension and continuation of the author's comments regarding the combining of criterion measures.

Taft, R., "Multiple Methods of Personality Assessment", in Psychological Bulletin, Vol. 56, 1959, p. 333 - 352.

One of the earliest articles which reviews and discusses various issues related to the use of multiple assessment techniques.

Tyler, Leona E., "Design for a Hopeful Psychology", in American Psychologist, Vol. 28, 1973, p. 1020 - 1029.

A global perspective of psychology is offered, part of which involves a critical look at the concept of psychic determinism and its implication for the science of psychology.

Ungerson, Bernard, "Assessment Centres - a Review of Research Findings", in Personnel Review, Vol. 3, 1974, p. 4 - 13.

A critical review of assessment centre research, leading to the conclusion that the most important unresolved issue involves establishing the relative validity of assessment centres in comparison with more traditional methods of management selection.

Sawyer, J., "Measurement and Prediction, Clinical and Statistical", in Psychological Bulletin, Vol. 66, 1966, p. 178-200.

Wernimont, Paul F., Management Assessment: Relationships Between Test Scores and Supervisor's Ratings on Job Performance Dimensions, Research Report No. 67, 3M Company, 1973.

This research study provides predictor-criterion data on several test measures used as part of an assessment centre program.

----- and Campbell, John P., "Signs, Samples, and Criteria", in Journal of Applied Psychology, Vol. 52, 1968, p. 372 - 376.

The authors present a behaviour consistency validation model which emphasizes the use of samples of behaviour rather than predispositions or signs of behaviour.

Wesman, Alexander G., "Intelligent Testing", in American Psychologist, Vol. 23, 1968, p. 264 - 274.

A general perspective is offered regarding the nature and testing of intelligence.

Wilson, John E. and Tatge, Walter A., "Assessment Centres: Further Assessment Needed?", in Personnel Psychology, Vol. 26, 1973, p. 172 - 179.

The first of several critical reviews of the assessment centre method emphasizing the lack of research evidence available to show the relative or comparative predictive validity of various assessment methods.

Wollowick, Herbert B., Personal communication, 1978.

A personal communication to clarify some of the methodological procedures used in the study noted below.

----- and McNamara, W.J., "Relationship of the Components of an Assessment Center to Management Success", in Journal of Applied Psychology, Vol. 53, 1969, p. 348 - 352.

The most frequently quoted study to illustrate the relative and incremental validity associated with situational and paper-and-pencil components of an assessment centre process.

Worbois, G.M., "Validation of Externally Developed Assessment Procedures for Identification of Supervisory Potential", in Personnel Psychology, Vol. 28, 1975, p. 77 - 91.

A concurrent validation study relating assessment centre performance to global criterion measures of job performance.

APPENDIX I

FREQUENCY AND PERCENTAGE DISTRIBUTIONS
FOR RELEVANT SUBJECT VARIABLES

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE VARIABLE OF LANGUAGE OF ASSESSMENT

Language of Assessment	Frequency	Percentage
English	237	80%
French	60	20%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE VARIABLE OF SEX

Sex	Frequency	Percentage
Male	249	84%
Female	48	16%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE VARIABLE OF AGE

Age	Frequency	Percentage
26-27	14	6%
28-29	37	12%
30-31	38	13%
32-33	46	15%
34-35	49	16%
36-37	28	9%
38-39	29	10%
40-41	19	7%
42-43	18	6%
44-45	12	4%
46-47	2	1%
48-49	2	1%
TOTAL	294*	100%

* data missing for 3 subjects

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE VARIABLE OF EDUCATIONAL LEVEL

Educational Level	Frequency	Percentage
High School or Less	20	6%
University or College Degree or Diploma	176	60%
Post-graduate Degree	98	34%
TOTAL	294 *	100%

* data missing for 3 subjects

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE VARIABLE OF OCCUPATIONAL CATEGORY

Occupational Category	Frequency	Percentage
Administrative	223	75%
Scientific and Professional	74	25%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE VARIABLE OF CLASSIFICATION GROUP

Classification Group	Frequency	Percentage
Program Managers	61	21%
Personnel	52	17%
Administration	37	13%
Engineering	21	7%
Commerce	16	5%
Finance	15	5%
Economics	15	5%
Organization and Methods	11	4%
Education	11	4%
Information Services	10	3%
Other*	48	16%
TOTAL	297	100%

* Computer Systems, Auditing, Scientific Research, Purchasing, Welfare Programs, Architecture, Meteorology, and so forth.

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE VARIABLE OF TOTAL YEARS WORKED

Total Years Worked	Frequency	Percentage
1 - 3	4	1%
4 - 6	35	12%
7 - 9	70	24%
10 - 12	55	19%
13 - 15	36	12%
16 - 18	34	11%
19 - 21	25	9%
22 - 24	23	8%
25 - 27	7	3%
28 - 30	4	1%
TOTAL	293*	100%

* data missing for 4 subjects

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE VARIABLE OF TOTAL YEARS WITHIN PUBLIC SERVICE

Total Years Within Public Service	Frequency	Percentage
1 - 3	50	17%
4 - 6	80	27%
7 - 9	90	31%
10 - 12	29	10%
13 - 15	16	6%
16 - 18	13	4%
19 - 21	10	3%
22 - 24	3	1%
25 - 27	2	1%
TOTAL	293*	100%

* data missing for 4 subjects

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE VARIABLE OF YEARS IN PRESENT POSITION

Years in Present Position	Frequency	Percentage
0 - 1	114	39%
2 - 3	133	45%
4 - 5	26	8%
6 - 7	15	5%
8 - 9	1	1%
10 or more	4	2%
TOTAL	293*	100%

* data missing for 4 subjects

APPENDIX II

DESCRIPTION OF THE SUB-TESTS
OF THE GENERAL INTELLIGENCE TEST

DESCRIPTION OF THE SUB-TESTS
OF THE GENERAL INTELLIGENCE TEST

The General Intelligence Test is subdivided into six parts or sub-tests. These are arranged in the following order:

Sub-Test I: General Information

This sub-test samples information of a factual nature which the candidate has gathered from his environment. It is information which an average individual with average opportunity would be able to acquire. Results on the sub-test give an indication of the candidate's educational environment and cultural predilections. Long-term memory is also tested here.

Example: The capital of Canada is:

- | | | |
|--------------|--------------|--------------|
| (1) Montreal | (2) Toronto | (3) Kingston |
| (4) Ottawa | (5) Victoria | |

Sub-Test II: Similarities

This sub-test is a verbal abstract reasoning test. Results from this sub-test indicate a candidate's ability to form concepts. Concept formation involves recognizing that certain objects or events belong together, and it provides the structure for organizing new experience into patterns of conceptual coherence. There are three main levels of concept formation which are qualitatively different - concrete, functional and abstract-conceptual. Because it requires the highest level of intelligence, it is the abstract-conceptual level of concept formation which this sub-test is designed to measure.

Example: Select the pair made up of words expressing exactly the same concept. There is only one such pair in the question

- (1) apple - automobile
- (2) orange - apple
- (3) automobile - cloud
- (4) charity - automobile
- (5) cloud - charity

Sub-Test III: Arithmetic Problem Solving

This sub-test is a non-verbal reasoning test involving one of the purest forms of abstract thinking - arithmetic. Arithmetic requires concentration to retrieve information about facts and relationships in the test patterns. The problems in the sub-test involve commonplace situations and practical calculations. Items involve the four basic types of calculation: addition, subtraction, multiplication and division.

Example: Without doing any written calculations, you are to solve each problem and indicate which answer you think is correct.

Fourteen is two more than half a certain number.
What number is two more than that certain number.

- (1) 26 (2) 28 (3) 25 (4) 24 (5) 29

Sub-Test IV: Vocabulary

This sub-test measures the extent of the candidate's vocabulary. The words used are reasonably common and would be encountered by most adults with varying degrees of frequency.

Example: From among the five choices given, select the word or phrase which best defines the word at the beginning of the question.

DEBRIS

- (1) covering
- (2) scattered fragments
- (3) gravel
- (4) garbage
- (5) to crack or split

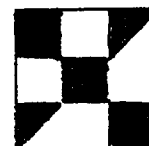
Sub-Test V: Block Design

This sub-test is an abstract reasoning test which uses figural patterns. It measures to some extent the candidate's perceptual organization ability. Essentially a non-verbal test, it is composed of geometrical patterns which must be analyzed and resynthesized visually to produce an answer.

Example: In each question the candidate is to decide how many times the type of block shown at the beginning of the question appears on a slightly reduced scale in the nine-block design on the right of the question.



appears in the design on the right exactly (1) (2) (3) (4) (5) times.

Sub-Test VI: Memory

This sub-test measures short-term memory, that is, how much unfamiliar new material is remembered after a twelve-minute period during which a different mental activity was carried out.

Example: In the memory sub-test, the candidate is presented with five paragraphs which he is to try and memorize as much as possible within 12 minutes. He is later given the same text which will contain numbered blank spaces. Each numbered blank space represents a question.

Original Text: "The Fifth Army shared in the French offensive on July 18."

Modified Text: "The (1) Army shared in the French
offensive on July 18."

1. (1) Fourth (2) Fifth (3) Second (4) First
(5) Third

APPENDIX III

DESCRIPTION OF THE DIMENSIONS OF
THE PERSONALITY RESEARCH FORM

DESCRIPTION OF THE DIMENSIONS OF
THE PERSONALITY RESEARCH FORM

The Personality Research Form assesses twenty-two dimensions of personality. These can be subsumed under seven superordinate categories. For each dimension, there is a description of a high scorer. As well, each dimension has associated with it a series of defining trait adjectives.

1. Measures of Orientation toward Work and Play

- a. Achievement: Aspires to accomplish difficult tasks; maintains high standards and is willing to work toward distant goals; responds positively to competition; willing to put forth effort to attain excellence.

Striving, accomplishing, capable, purposeful, attaining, industrious, achieving, aspiring, enterprising, self-improving, productive, driving, ambitious, resourceful, competitive.

- b. Endurance: Willing to work long hours; doesn't give up quickly on a problem; persevering, even in the face of great difficulty; patient and unrelenting in his work habits.

Persistent, determined, steadfast, enduring, unfaltering, persevering, unremitting, relentless, tireless, dogged, energetic, has stamina, sturdy, zealous, durable.

- c. Play: Does many things "just for fun"; spends a good deal of time participating in games, sports, social activities, and other amusements; enjoys jokes and funny stories; maintains a light-hearted, easy-going attitude toward life.

Playful, jovial, jolly, pleasure-seeking, merry, laughter-loving, joking, frivolous, prankish, sportive, mirthful, fun-loving, gleeful, carefree, blithe.

2. Measures of Degree of Ascendency

- a. Abasement: Shows a high degree of humility; accepts blame and criticism even when not deserved; exposes himself to situations where he is in an inferior position; tends to be self-effacing.

Meek, self-accusing, self-blaming, obsequious, self-belittling, surrendering, resigned self-critical, humble, apologizing, subservient, obedient, yielding, deferential, self-subordinating.

- b. Dominance: Attempts to control his environment, and to influence or direct other people; expresses opinions forcefully; enjoys the role of leader and may assume it spontaneously.

Governing, controlling, commanding, domineering, influential, persuasive, forceful, ascendant, leading, directing, dominant, assertive, authoritative, powerful, supervising.

3. Measures of Orientation Towards Direction from Other People

- a. Autonomy: Tries to break away from restraints, confinement, or restrictions of any kind; enjoys being unattached, free, not tied to people, places, or obligations; may be rebellious when faced with restraints.

Unmanageable, free, self-reliant, independent, autonomous, rebellious, unconstrained, individualistic, ungovernable, self-determined, non-conforming, uncompliant, undominated, resistant, lone-wolf.

- b. Succourance: Frequently seeks the sympathy, protection, love, advice, and reassurance of other people; may feel insecure or helpless without such support; confides difficulties readily to a receptive person.

Trusting, ingratiating, dependent, entreating, appealing for help, seeks support, wants advice, helpless, confiding, needs protection, requesting, craves affection, pleading, help seeking, defensless.

4. Measures of Degree and Quality of Interpersonal Orientation

- a. Affiliation: Enjoys being with friends and people in general; accepts people readily; makes efforts to win friendships and maintain associations with people.

Neighbourly, loyal, warm, amicable, good-natured, friendly, companionable, genial, affable, cooperative, gregarious, hospitable, sociable, affiliative, good-willed.

- b. Aggression: Enjoys combat and argument; easily annoyed; sometimes willing to hurt people to get his way; may seek to "get even" with people whom he perceives as having harmed him.

Aggressive, quarrelsome, irritable, argumentative, threatening, attacking, antagonistic, pushy, hot-tempered, easily-angered, hostile, revengeful, belligerent, blunt, retaliative.

- c. Defence: Readily suspects that people mean him harm or are against him; ready to defend himself at all times; takes offense easily; does not accept criticism readily.

Self-protective, justifying, denying, defensive, self-condoning, suspicious, secretive, has a "chip on the shoulder", resists inquiries, protesting, wary, self-excusing, rationalizing, guarded, touchy.

- d. Exhibition: Wants to be the center of attention; enjoys having an audience; engages in behavior which wins the notice of others, may enjoy being dramatic or witty.

Colorful, entertaining, unusual, spellbinding, exhibitionistic, conspicuous, noticeable, expressive, ostentatious, immodest, demonstrative, flashy, dramatic, pretentious, showy.

- e. Nurturance: Gives sympathy and comfort; assists others whenever possible, interested in caring for children, the disabled, or the infirm; offers a "helping hand" to those in need; readily performs favors for others.

Sympathetic, paternal, helpful, benevolent, encouraging, caring, protective, comforting, maternal, supporting, aiding, ministering, consoling, charitable, assisting.

- e. Social Recognition: Desires to be held in high esteem by acquaintances; concerned about reputation and what other people think of him; works for the approval and recognition of others.

Approval seeking, proper, well-behaved, seeks recognition, courteous, makes good impression, seeks respectability, accommodating, socially proper, seeks admiration, obliging, agreeable, socially sensitive, desirous of credit, behaves appropriately.

5. Measures of Impulse Expression and Control

- a. Change: Likes new and different experiences; dislikes routine and avoids it; may readily change opinions or values in different circumstances; adapts readily to changes in environment.

Inconsistent, fickle, flexible, unpredictable, wavering, mutable, adaptable, changeable, irregular, variable, capricious, innovative, fighty, vacillating, inconstant.

- b. Cognitive Structure: Does not like ambiguity or uncertainty in information; wants all questions answered completely; desires to make decisions based upon definite knowledge, rather than upon guesses or probabilities.

Precise, exacting, definite, seeks certainty, meticulous, perfectionistic, clarifying, explicit, accurate, rigorous, literal, avoids ambiguity, defining, rigid, needs structure.

- c. Harmavoidance: Does not enjoy exciting activities, especially if danger is involved; avoids risk of bodily harm; seeks to maximize personal safety.

Fearful, withdraws from danger, self-protecting, pain-avoidant, careful, cautious, seeks safety, timorous, apprehensive, precautionary, unadventurous, avoids risks, attentive to danger, stays out of harm's way, vigilant.

- d. Impulsivity: Tends to act on the "spur of the moment" and without deliberation; gives vent readily to feelings and wishes; speaks freely; may be volatile in emotional expression.

Hasty, rash, uninhibited, spontaneous, reckless, irrepresible, quick-thinking, mecurial, impatient, incautious, hurried, impulsive, foolhardy, excitable, impetuous.

- e. Order: Concerned with keeping personal effects and surroundings neat and organized; dislikes clutter, confusion, lack of organization; interested in developing methods for keeping materials methodically organized.

Neat, organized, tidy, systematic, well-ordered, disciplined, prompt, consistent, orderly, clean, methodical, scheduled, planful, unvarying, deliberate.

6. Measures of Intellectual and Aesthetic Orientations

- a. Sentience: Notices smells, sounds, sights, tastes, and the way things feel; remembers these sensations and believes that they are an important part of life; is sensitive to many forms of experience; may maintain an essentially hedonistic or aesthetic view of life.

Aesthetic, enjoys physical sensations, observant, earthy, aware, notices environment, feeling, sensitive, sensuous, open to experience, perceptive, responsive, noticing, discriminating, alive to impressions.

- b. Understanding: Wants to understand many areas of knowledge; values synthesis of ideas, verifiable generalization, logical thought, particularly when directed at satisfying intellectual curiosity.

Inquiring, curious, analytical, exploring, intellectual, reflective, incisive, investigative, probing, logical, scrutinizing, theoretical, astute, rational, inquisitive.

7. Measures of Test-taking Attitudes and Validity

- a. Desirability: Describes self in terms judged as desirable; consciously or unconsciously, accurately or inaccurately, presents favorable picture of self in responses to personality statements.
- b. Infrequency: Responds in implausible or pseudo-random manner, possibly due to carelessness, poor comprehension, passive non-compliance, confusion, or gross deviation.

APPENDIX IV

EQUIVALENCY CONVERSION TABLE FOR
VARIOUS CLASSIFICATION GROUPS

EQUIVALENCY CONVERSION TABLE FOR
VARIOUS CLASSIFICATION GROUPS

The equivalency table noted on the following page equates the position levels within the organization for the different classification groups. The following list gives the letter code for each of the twenty-eight classification groups represented in the sample used for this study.

PM - Program Managers	WP - Welfare Programs
PE - Personnel	FS - Foreign Service
AS - Administration	TR - Translation
EN - Engineering	AR - Architecture
CO - Commerce	MT - Meteorology
FI - Finance	CH - Chemistry
ES - Economics	AG - Agriculture
OM - Organization and Methods	FO - Forestry
ED - Education	LA - Law
IS - Information Services	LS - Library Science
CS - Computer Systems	MA - Mathematics
SE - Scientific Research	PS - Psychology
AU - Auditing	SG - Scientific Regulation
PG - Purchasing	SW - Social Work

(October 20, 1977)

100-1001

75-5 - 5X-3

10.5 9X-2

11.5 92-3

12.5 3X-4

APPENDIX V

FREQUENCY AND PERCENTAGE DISTRIBUTIONS FOR
THE PREDICTOR AND CRITERION MEASURES

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE PAPER & PENCIL TEST MEASURE OF GENERAL INTELLIGENCE

General Intelligence	Frequency	Percentage
31 - 35	4	2%
36 - 40	6	2%
41 - 45	9	3%
46 - 50	9	3%
51 - 55	23	8%
56 - 60	30	9%
61 - 65	38	13%
66 - 70	51	17%
71 - 75	55	18%
76 - 80	38	13%
81 - 85	23	8%
86 - 90	10	3%
91 - 95	1	1%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE PAPER & PENCIL TEST MEASURE OF ADMINISTRATIVE JUDGEMENT

Administrative Judgement	Frequency	Percentage
10 - 14	1	1%
15 - 19	14	5%
20 - 24	71	24%
25 - 29	121	40%
30 - 34	73	24%
35 - 39	16	5%
40 - 44	1	1%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE PAPER & PENCIL TEST MEASURE OF ACHIEVEMENT

Achievement	Frequency	Percentage
7 - 8	1	1%
9 - 10	9	3%
11 - 12	20	7%
13 - 14	44	15%
15 - 16	91	30%
17 - 18	90	30%
19 - 20	42	14%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE PAPER & PENCIL TEST MEASURE OF ENDURANCE

Endurance	Frequency	Percentage
3 - 4	1	1%
5 - 6	4	2%
7 - 8	10	3%
9 - 10	19	6%
11 - 12	51	17%
13 - 14	61	20%
15 - 16	78	26%
17 - 18	58	20%
19 - 20	15	5%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE PAPER & PENCIL TEST MEASURE OF DOMINANCE

Dominance	Frequency	Percentage
3 - 4	4	1%
5 - 6	3	1%
7 - 8	8	3%
9 - 10	25	8%
11 - 12	54	18%
13 - 14	65	22%
15 - 16	75	25%
17 - 18	55	19%
19 - 20	8	3%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE PAPER & PENCIL TEST MEASURE OF ABASEMENT

Abasement	Frequency	Percentage
1 - 2	39	13%
3 - 4	57	19%
5 - 6	91	31%
7 - 8	63	21%
9 - 10	32	11%
11 - 12	12	4%
13 - 14	3	1%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE PAPER & PENCIL TEST MEASURE OF AUTONOMY

Autonomy	Frequency	Percentage
1 - 2	3	1%
3 - 4	23	8%
5 - 6	49	16%
7 - 8	76	26%
9 - 10	84	28%
11 - 12	42	14%
13 - 14	15	5%
15 - 16	5	2%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE PAPER & PENCIL TEST MEASURE OF SUCCOURANCE

Succourance	Frequency	Percentage
1 - 2	9	3%
3 - 4	47	16%
5 - 6	67	23%
7 - 8	81	27%
9 - 10	47	16%
11 - 12	31	10%
13 - 14	9	3%
15 - 16	6	2%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE PAPER & PENCIL TEST MEASURE OF AFFILIATION

Affiliation	Frequency	Percentage
3 - 4	1	1%
5 - 6	2	1%
7 - 8	10	3%
9 - 10	29	10%
11 - 12	37	12%
13 - 14	54	18%
15 - 16	87	29%
17 - 18	63	21%
19 - 20	14	5%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE SITUATIONAL TEST MEASURE OF STRESS TOLERANCE

Stress Tolerance	Frequency	Percentage
1	3	1%
2	29	9%
3	65	22%
4	89	30%
5	70	24%
6	40	13%
7	1	1%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE SITUATIONAL TEST MEASURE OF MOTIVATION

Motivation	Frequency	Percentage
1	5	2%
2	20	7%
3	51	17%
4	80	27%
5	77	26%
6	59	19%
7	5	2%
TOTAL	297	100%

FREQUENCY AND PERCENTILE DISTRIBUTION
FOR THE SITUATIONAL TEST MEASURE OF INDEPENDENCE

Motivation	Frequency	Percentage
1	9	3%
2	39	13%
3	70	24%
4	78	26%
5	66	22%
6	32	11%
7	3	1%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE SITUATIONAL TEST MEASURE OF LEADERSHIP

Leadership	Frequency	Percentage
1	12	4%
2	49	16%
3	73	25%
4	68	23%
5	65	22%
6	27	9%
7	3	1%
Total	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE SITUATIONAL TEST MEASURE OF INTERPERSONAL RELATIONS

Interpersonal Relations	Frequency	Percentage
1	3	1%
2	25	8%
3	65	22%
4	95	32%
5	64	22%
6	38	13%
7	7	2%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE SITUATIONAL TEST MEASURE OF ANALYSIS AND SYNTHESIS

Analysis and Synthesis	Frequency	Percentage
1	11	4%
2	34	11%
3	67	23%
4	78	26%
5	69	23%
6	34	11%
7	4	2%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE SITUATIONAL TEST MEASURE OF ORAL COMMUNICATION

Oral Communication	Frequency	Percentage
1	2	1%
2	24	8%
3	55	19%
4	110	37%
5	57	19%
6	47	15%
7	2	1%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE SITUATIONAL TEST MEASURE OF QUALITY OF JUDGEMENT

Quality of Judgement	Frequency	Percentage
1	11	4%
2	39	13%
3	100	33%
4	82	28%
5	43	14%
6	20	7%
7	2	1%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE SITUATIONAL TEST MEASURE OF INITIATING ACTIVITIES

Initiating Activities	Frequency	Percentage
10 - 14	11	4%
15 - 19	32	10%
20 - 24	50	16%
25 - 29	64	22%
30 - 34	65	22%
35 - 39	35	11%
40 - 44	23	8%
45 - 49	11	4%
50 - 54	5	2%
55 - 59	1	1%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE SITUATIONAL TEST MEASURE OF DECISION MAKING

Decision Making	Frequency	Percentage
46 - 50	10	3%
51 - 55	50	17%
56 - 60	110	37%
61 - 65	91	31%
66 - 70	34	11%
71 - 75	2	1%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE SITUATIONAL TEST MEASURE OF PROBLEM COMPREHENSION

Problem Comprehension	Frequency	Percentage
46 - 50	4	2%
51 - 55	49	16%
56 - 60	114	38%
61 - 65	90	30%
66 - 70	35	12%
71 - 75	5	2%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE SITUATIONAL TEST MEASURE OF DELEGATION

Delegation	Frequency	Percentage
47 - 48	7	2%
49 - 50	77	26%
51 - 52	161	54%
53 - 54	41	14%
55 - 56	11	4%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE SITUATIONAL TEST MEASURE OF PLANNING AND ORGANIZATION

Planning & Organization	Frequency	Percentage
1 - 5	4	2%
6 - 10	27	9%
11 - 15	45	15%
16 - 20	47	16%
21 - 25	44	15%
26 - 30	40	13%
31 - 35	41	14%
36 - 40	31	10%
41 - 45	6	2%
46 - 50	10	3%
51 - 55	2	1%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE SITUATIONAL TEST MEASURE OF PRODUCTIVITY

Productivity	Frequency	Percentage
31 - 40	5	2%
41 - 50	12	4%
51 - 60	31	10%
61 - 70	41	14%
71 - 80	72	24%
81 - 90	73	25%
91 - 100	44	15%
101 - 110	16	5%
111 - 120	3	1%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE CRITERION MEASURE OF RATE OF POSITION-LEVEL INCREASE

Rate of Position-Level Increase	Frequency	Percentage
2.00 - 2.49	1	1%
1.50 - 1.99	4	2%
1.00 - 1.49	27	8%
.50 - .99	73	24%
0 - .49	186	62%
-.50 - -.01	4	2%
-1.00 - -.51	2	1%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE CRITERION MEASURE OF RATE OF SALARY INCREASE

Rate of Salary Increase	Frequency	Percentage
25.0 - 29.9	4	2%
20.0 - 24.9	14	4%
15.0 - 19.9	19	6%
10.0 - 14.9	49	16%
5.0 - 9.9	95	32%
0 - 4.9	71	24%
-5.0 - -0.1	24	8%
-10.0 - -5.1	17	5%
-15.0 - -10.1	1	1%
-20.0 - -15.1	2	1%
-25.0 - 20.1	1	1%
TOTAL	297	100%

FREQUENCY AND PERCENTAGE DISTRIBUTION
FOR THE COMPOSITE CRITERION MEASURE

Composite	Frequency	Percentage
1.50 - 1.99	4	2%
1.00 - 1.49	19	6%
.50 - .99	34	11%
0 - .49	87	29%
-.50 - -.01	86	29%
-1.00 - -.51	42	14%
-1.50 - -1.01	18	6%
-2.00 - -1.51	6	2%
-2.50 - -2.01	1	1%
TOTAL	297	100%

APPENDIX VI

THE MULTIPLE R'S AND ASSOCIATED REGRESSION WEIGHTS
OBTAINED WHEN THE PAPER-AND-PENCIL TEST MEASURES
AND THE SITUATIONAL TEST MEASURES ARE
RELATED TO THE CRITERION MEASURES

THE MULTIPLE R's AND ASSOCIATED REGRESSION WEIGHTS
OBTAINED WHEN THE NINE PAPER-AND-PENCIL TEST MEASURES
ARE RELATED TO THE CRITERION MEASURE OF RATE OF POSITION-LEVEL INCREASE

Paper-and-Pencil Test Measure	Multiple R	Regression Weight
General Intelligence	0.011	0.166
Administrative Judgement	0.085	0.545
Achievement	0.085	0.322
Endurance	0.096	0.265
Dominance	0.100	-0.483
Abasement	0.130	-0.156
Autonomy	0.138	-0.140
Succourance	0.194	-0.275
Affiliation	0.246	0.231
(Constant)	-	0.243

THE MULTIPLE R's AND ASSOCIATED REGRESSION WEIGHTS
OBTAINED WHEN THE NINE PAPER-AND-PENCIL TEST MEASURES
ARE RELATED TO THE CRITERION MEASURE OF RATE OF SALARY INCREASE

Paper-and-Pencil Test Measure	Multiple R	Regression Weight
General Intelligence	0.025	-0.198
Administrative Judgement	0.056	0.681
Achievement	0.059	0.426
Endurance	0.081	-0.248
Dominance	0.129	0.181
Abasement	0.156	-0.250
Autonomy	0.167	-0.283
Succourance	0.190	-0.302
Affiliation	0.196	0.137
(Constant)	-	10.664

THE MULTIPLE R'S AND ASSOCIATED REGRESSION WEIGHTS
OBTAINED WHEN THE NINE PAPER-AND-PENCIL TEST MEASURES
ARE RELATED TO THE COMPOSITE CRITERION MEASURE

Paper-and-Pencil Test Measure	Multiple R	Regression Weight
General Intelligence	0.011	0.374
Administrative Judgement	0.066	0.762
Achievement	0.066	0.582
Endurance	0.068	-0.121
Dominance	0.090	0.137
Abasement	0.140	-0.283
Autonomy	0.142	-0.212
Succourance	0.195	-0.410
Affiliation	0.216	0.228
(Constant)	-	0.100

THE MULTIPLE R's AND ASSOCIATED REGRESSION WEIGHTS
OBTAINED WHEN THE FOURTEEN SITUATIONAL TEST MEASURES ARE
RELATED TO THE CRITERION MEASURE OF RATE OF POSITION-LEVEL INCREASE

Situational Test Measure	Multiple R	Regression Weight
Stress Tolerance	0.040	-0.056
Motivation	0.175	0.021
Independence	0.192	0.010
Leadership	0.200	0.030
Interpersonal Relations	0.200	-0.010
Analysis & Synthesis	0.220	0.080
Oral Communication	0.221	-0.020
Quality of Judgement	0.222	-0.027
Initiating Activities	0.236	-0.001
Decision Making	0.238	0.010
Problem Comprehension	0.275	-0.020
Delegation	0.275	-0.001
Planning & Organization	0.277	0.001
Productivity	0.300	0.010
(Constant)	-	0.734

THE MULTIPLE R's AND ASSOCIATED REGRESSION WEIGHTS
OBTAINED WHEN THE FOURTEEN SITUATIONAL TEST MEASURES ARE
RELATED TO THE CRITERION MEASURE OF RATE OF SALARY INCREASE

Situational Test Measure	Multiple R	Regression Weight
Stress Tolerance	0.053	-1.150
Motivation	0.053	-0.985
Independence	0.113	0.544
Leadership	0.140	0.390
Interpersonal Relations	0.162	0.580
Analysis & Synthesis	0.170	0.434
Oral Communication	0.171	-0.222
Quality of Judgement	0.175	0.271
Initiating Activities	0.180	-0.010
Decision Making	0.190	0.100
Problem Comprehension	0.194	-0.110
Delegation	0.195	0.105
Planning & Organization	0.196	0.020
Productivity	0.199	0.028
(Constant)	-	0.913

THE MULTIPLE R's AND ASSOCIATED REGRESSION WEIGHTS
OBTAINED WHEN THE FOURTEEN SITUATIONAL TEST MEASURES ARE
RELATED TO THE COMPOSITE CRITERION MEASURE

Situational Test Measure	Multiple R	Regression Weight
Stress Tolerance	0.020	-0.122
Motivation	0.121	-0.027
Independence	0.145	0.005
Leadership	0.178	0.068
Interpersonal Relations	0.183	0.026
Analysis & Synthesis	0.225	0.158
Oral Communication	0.233	-0.042
Quality of Judgement	0.236	-0.054
Initiating Activities	0.240	-0.004
Decision Making	0.248	0.012
Problem Comprehension	0.269	-0.023
Delegation	0.270	0.008
Planning & Organization	0.272	0.002
Productivity	0.282	0.006
(Constant)	-	-0.245

APPENDIX VII

ABSTRACT OF

A TEST OF THE RELATIVE AND INCREMENTAL PREDICTIVE
VALIDITY ASSOCIATED WITH A SET OF PAPER-AND-PENCIL
TEST MEASURES AND A SET OF SITUATIONAL TEST
MEASURES WITHIN AN ASSESSMENT CENTRE

ABSTRACT OF

A TEST OF THE RELATIVE AND INCREMENTAL PREDICTIVE
VALIDITY ASSOCIATED WITH A SET OF PAPER-AND-PENCIL
TEST MEASURES AND A SET OF SITUATIONAL TEST
MEASURES WITHIN AN ASSESSMENT CENTRE ¹

A sample of two hundred and ninety-seven subjects from lower middle to upper middle levels of management within the Canadian Federal Government participated in an assessment centre program designed to facilitate the identification of potential for effective functioning at executive levels of management. During the assessment program, performance was assessed through a series of situational tests and a battery of psychometric paper-and-pencil tests. The general progression achieved by these participants approximately one and one-half to four years following assessment was evaluated by means of criterion measures related to position-level and salary, as well as a composite measure combining both of these criterion variables. The research was primarily designed to compare the predictive validity of a set of paper-and-pencil test measures with the validity associated with a set of situational test measures. As part of this effort, the research also focussed on the incremental validity which could ensue from the combination of both sets of predictors. The results of the investigation indicated

1 Leonard W. Slivinski, doctoral thesis presented to the School of Psychology of the University of Ottawa, Ontario, August, 1978, VIII - 167 p.

that there was no significant difference in the amount of predictive variance accounted for by the paper-and-pencil versus situational testing methods. It was determined, however, that both sets of predictors added in an incremental fashion in the overall prediction of progression within the organization. The existence of this relationship was to some extent dependent upon the criterion measure considered.