

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

AN EVALUATION OF SELF-ASSESSMENT
FOR PERSONNEL SELECTION IN NEW ZEALAND

A thesis presented in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy
in Psychology at Massey University

DAVID IAN GEORGE

1988

ABSTRACT

Personnel selection decisions in New Zealand and other countries are commonly made by the employer. Although applicants have as much, if not more to gain or lose from a selection decision, they are rarely allowed to contribute to the decision making process as an equal party to a contract. Psychological theory and research shows that while there are a number of factors which moderate the accuracy of self-judgments, individuals are capable of assessing their own abilities and performance, particularly in work and work related settings. To examine the acceptability and validity of self-assessment for New Zealand institutional personnel selection, 69 organisations were asked to include self-assessment in their selection procedures. Self-assessment was implemented in 14 of the 69 organisations. These organisations had agreed to cooperate and expected sufficiently large recruitment levels to make the research feasible. Of the 14 settings, four completed the project providing sufficient data for an evaluation of validity of self-assessment. In addition to examining the relationship of self-assessment with work performance criteria, a number of possible moderator variables were examined. The results of this research showed firstly, that few New Zealand practitioners considered self-assessment could be a useful part of their selection procedures, secondly, the validity offered by self-assessment was poor, and thirdly, no moderator effects were detected consistently across the studies conducted in different settings. It appears that until practitioners in New Zealand and elsewhere allow candidates to collaborate in selection decisions, the demand characteristics of the traditional selection situation will hamper accurate self-assessment.

ACKNOWLEDGEMENTS

I am exceptionally grateful to my supervisor Mike Smith for the support, friendship, and encouragement he gave me through much of my university study, especially the research and writing involved in this dissertation. I would describe Mike as an exceptionally effective motivator. I found it was always more attractive to meet a deadline than to put any effort into explaining why there was a lack of progress. His commitment to completion of my thesis was particularly appreciated.

I realise that my immediate family, Alison and our daughters Jenny and Sharon, have made a substantial investment in me. It is generally not possible for one family member to make what is often a selfish commitment to academic study. I am extremely grateful that they have allowed me to forgo many family responsibilities to meet the requirements of my research. I am also indebted to our parents, Hazel, Harry and Judith, for their support whenever it was needed.

John Podd, my second supervisor, is a generous friend who rarely says no. I am thankful for the many times when he had only moments to spare but would spend hours talking over problems and offering advice. He also deserves a special commendation as a "spotter" of split infinitives.

The staff in the Department of Psychology have been a supportive and caring group of professionals. For me their company has been both productive and enjoyable in a range of sporting, social, and academic activities. I particularly appreciate the time willingly given by John Spicer to advise and discuss the statistical knots I encountered and the support and understanding offered by Professor George Shouksmith as head of department.

I am very thankful for the cooperation and friendship of those practitioners who were involved in the project. Thank you Mike Wilson, John Yeabsley, Francis Blyth, Shona Butterfield, Helen Clelland, Lorraine Williams, and Denis Schofield. Your efforts were largely responsible for the survival of the different studies, often in the face of numerous other demands on your time. I am also appreciative of the time given by numerous individuals who as applicants or employees were part of the research.

I could not have considered a research project of this size without the financial support given to me by the University Grants Committee (U.G.C. Scholarship, Shirtcliffe Fellowship) and BP (NZ) Ltd. (BP Scholarship). In addition, access to administrative and computer resources given by the Psychology Department and Massey University was invaluable. Thank you for your help.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGEMENTS	iii
CHAPTER ONE	
AN OVERVIEW	1
CHAPTER TWO	
CURRENT PERSONNEL SELECTION AND PRACTICE	6
Current Issues	6
Predictor Research and Practice	11
Psychological Tests	11
Interviews	15
Application Form Information	17
References, Letters of Recommendation	19
Work Sample Tests	20
Assessment Centres	21
Other Predictors	23
Measures of Work Performance	26
Personnel Selection in New Zealand	29
CHAPTER THREE	
THE PSYCHOLOGY OF SELF.	35
Structure to Process: Strategic Self-Presentation	37
Mechanism to Agency: Reciprocal Determinism and Self-Efficacy	42
Self-Knowledge to Social Construction: Social Comparison	46

Self-Assessment and the Self	48
--	----

CHAPTER FOUR

THE APPLICATION OF SELF-ASSESSMENT	54
Empirical Research Using Self-Assessment	54
Self-Assessment in Educational Research	57
Self-Assessment in Vocational Guidance	59
The Validity of Self-Assessment in Work and Work-Related Situations	61
Factors Affecting Self-Assessment in Work Situations	72
Approaches to Improving Self-Assessment Validity	77

CHAPTER FIVE

AN EVALUATION OF SELF-ASSESSMENT IN PERSONNEL SELECTION	81
The Acceptability of Self-Assessment in Institutional Personnel Selection	82
The New Zealand Context	83
Aims of the Research Project	89
The Acceptability of Self-Assessment in Personnel Selection	90
Method.	91
Settings	91
Procedure	92
Results.	93
Discussion	101

CHAPTER SIX

EVALUATION OF SELF-ASSESSMENT VALIDITY IN AN ELECTRONICS	
MANUFACTURING COMPANY	105
Method.	105
Subjects.	105

Procedure	105
Results	117
Discussion	120

CHAPTER SEVEN

EVALUATION OF SELF-ASSESSMENT VALIDITY IN A FOOD PROCESSING

COMPANY	123
Study A	124
Method	124
Subjects	124
Procedure	124
Results	130
Discussion	136
Study B	143
Method	143
Subjects	143
Procedure	143
Results	147
Discussion	154

CHAPTER EIGHT

EVALUATION OF SELF-ASSESSMENT VALIDITY IN A POLYTECHNIC NURSING

COURSE	159
Study A	160
Method	160
Subjects	160
Procedure	160
Results	163

Discussion	164
Study B	166
Method.	167
Subjects.	167
Procedure	167
Results.	170
Discussion	172

CHAPTER NINE

EVALUATION OF SELF-ASSESSMENT VALIDITY IN A GOVERNMENT

DEPARTMENT.	176
Study A	176
Method.	176
Subjects.	176
Procedure	177
Results.	180
Discussion	181
Study B	183
Method.	183
Subjects.	183
Procedure	184
Results.	188
Discussion	194

CHAPTER TEN

SELF-ASSESSMENT IN NEW ZEALAND PERSONNEL SELECTION 198

The Acceptability of Self-Assessment in New Zealand. 199

The Validity of Self-Assessment in New Zealand 201

Moderator Effects on the Validity of Self-Assessment	204
Possibilities for Further Research and Practice in New Zealand	206
REFERENCE LIST	209
APPENDIX 1: LETTER SENT REQUESTING COOPERATION IN RESEARCH PROJECT	253
APPENDIX 2: SELF-ASSESSMENT REQUEST USED IN ELECTRONICS MANUFACTURING COMPANY	254
APPENDIX 3: EMPLOYEE APPRAISAL FORMS USED IN ELECTRONICS MANUFACTURING COMPANY	257
APPENDIX 4: INTERVIEWERS GUIDE AND GENERAL KNOWLEDGE TEST USED IN ELECTRONICS MANUFACTURING COMPANY	259
APPENDIX 5: SUB-GROUP CORRELATIONS FROM ELECTRONIC ASSEMBLER SAMPLE	262
APPENDIX 6: SELF-ASSESSMENT REQUEST USED IN THE FOOD PROCESSING COMPANY (STUDY A)	264
APPENDIX 7: SELF-ASSESSMENT REQUEST USED IN THE FOOD PROCESSING COMPANY (STUDY B)	267
APPENDIX 8: SELF-ASSESSMENT REQUEST USED IN THE POLYTECHNIC NURSING COURSE (STUDY A)	269

**APPENDIX 9: SELF-ASSESSMENT REQUEST USED IN THE POLYTECHNIC
NURSING COURSE (STUDY B) 271**

**APPENDIX 10: SELF-ASSESSMENT REQUEST USED IN THE GOVERNMENT
DEPARTMENT (STUDY A) 273**

**APPENDIX 11: DESCRIPTION OF GOVERNMENT DEPARTMENT ASSESSMENT
CENTRE COMPONENTS 276**

**APPENDIX 12: SELF-ASSESSMENT REQUEST USED IN THE GOVERNMENT
DEPARTMENT (STUDY B) 279**

**APPENDIX 13: SUB-GROUP CORRELATIONS FROM THE DESK OFFICER SAMPLE
(GOVERNMENT DEPARTMENT STUDY B) 282**

LIST OF TABLES

Table 3.1

<u>Taxonomy of Self-Presentation Strategies Classified Primarily by Attribution Sought</u>	39
--	----

Table 4.1

<u>Application of Self-assessment as a Predictor in Work and Work Related Situations</u>	63
--	----

Table 4.2

<u>Sub-Groupings of Self-Assessment Correlations</u>	70
--	----

Table 5.1

<u>Sample of Organisations Classified Into Major Industry Groups.</u>	92
---	----

Table 5.2

<u>Organisations in Which Self-Assessment Research was Conducted</u>	95
--	----

Table 5.3

<u>Organisation Settings, Samples, Selection Procedures and Criteria Used in Analysis.</u>	100
--	-----

Table 7.1

<u>Correlations Between Self-Assessment and PERFORM for Four Age Quartile Groups</u>	132
--	-----

Table 7.2

Correlations Between Self-Assessment and Absence for Four Age
Quartile Groups 136

Table 7.3

Correlations Between Appraised Performance, Turnover, and Absence
Criteria (duration and frequency) 147

Table 7.4

Correlations Between SAJOB and TURN for Five Levels of Secondary
School Qualification 151

Table 8.1

Inter-rater Reliability Coefficients 163

Table 8.2

R² Values on First and Last Step of Moderated Regression Analysis 171

Table 9.1

Predictive Relationships Demonstrated Between Self-Assessment
and Work Criteria 181

Table 9.2

Characteristics of the Total Sample and the Employed Sample 184

LIST OF FIGURES

<u>Figure 7.1.</u>	Moderated Regression Results for Appraised Performance Criteria.	131
<u>Figure 7.2.</u>	Moderated Regression Results for Turnover Criteria.	134
<u>Figure 7.3.</u>	Moderated Regression Results for PERFORM (Appraised Job Performance).	149
<u>Figure 7.4.</u>	Moderated Regression Results for TURN Criteria.	150
<u>Figure 7.5.</u>	Moderated Regression Results for AB-DURATION (length of time absent from work without a telephoned reason)..	152
<u>Figure 7.6.</u>	Moderated Regression Results for AB-FREQUENCY (number of times absent from work without a telephoned reason).	153
<u>Figure 9.1.</u>	Self and Organisational Ratings of Job and Assessment Centre Performance Made Before and After an Assessment Centre.	190
<u>Figure 9.2.</u>	Relationships Between Self and Organisational Assessments of Job Performance Made During Selection and When Employed.	191
<u>Figure 9.3.</u>	Relationships Between Self-Assessments and Organisational Promotion Decisions Made During Selection and When Employed.	192

CHAPTER ONE

AN OVERVIEW

If there is any theme that characterises personnel selection research it is a focus on the quality of selection decisions. Selection decisions are traditionally made by the employer with the cooperation of the applicant who is required to supply information and submit to observation as requested. The applicant's role in the decision making process rarely extends beyond that of a data source. Applicants also have to make decisions. They have to decide whether or not they can meet job requirements, whether the job meets their requirements, and eventually, whether they will accept or reject a job offer. Practitioners and researchers have largely ignored the applicant's role in the employment process, especially the contribution an applicant may make to an institutional selection decision.

This research project considers an extension of the traditional role of an applicant beyond that of a data source, to providing judgments that can be used in a selection decision. While the potential of self-assessment has been recognised and used in some processes such as vocational guidance, this approach is rarely used in personnel selection. Conclusions regarding the acceptability of self-assessment have been made in other countries, notably the USA and United Kingdom, but comparable New Zealand research is lacking. The purpose of this investigation therefore is to assess the acceptability and validity of self-assessment for New Zealand institutional personnel selection.

Chapter Two begins by discussing current issues in personnel selection research and practice as background to the research situation. The profound effect of civil rights legislation on personnel selection is examined and illustrated by examples from various

parts of the world. The emphasis on improving the validity of various predictors is demonstrated in a range of research involving old and new techniques. Measures of work performance are discussed with particular attention to the less commonly used criteria. Personnel selection in New Zealand is examined suggesting little comprehensive information exists regarding the selection methods used. That which is available implies that New Zealand practitioners appear similar to their overseas counterparts in terms of their preference for traditional predictors, often of dubious validity. The role of the applicant as an assessor is rarely investigated.

An understanding of the problems and possibilities of self-assessment can be provided by an appreciation of the concept of self. Psychology as the study of human behaviour is vitally concerned with the self and is able to offer a theoretical and empirical background for this investigation. Chapter Three provides a psychological perspective of the self and explores the issues related to self-assessment. Recent trends in self research are discussed with particular emphasis on the active role the self has in the regulation of behaviour.

The use of self-assessment as a predictor is discussed in Chapter Four. Reference to two major reviews indicate that self-assessment may be an accurate predictor in a variety of applications. The validity of self-assessment in vocational guidance supports this contention. A discussion of self-assessment applied to educational achievement prediction maintains that the provision of such information appears to have few implications for the assessor so making the assessment less susceptible to bias. This contrasts with judgments made in selection contexts which can have a major impact on an assessor's working life.

The possibility that individuals could "tailor" self-assessments to gain desired outcomes emphasises the need to evaluate the utility of self-assessment in the situation in which it is to be used. Self-assessment research in work and work-related situations is reviewed and various individual and situational characteristics which may affect the validity of self-assessment are examined. It is shown that while self-assessment has potential as a predictor of work performance in selection situations, there are a number of factors which may moderate its validity. These include the way industry is organised and possibly the cultural environment. Efforts to improve the validity of self-assessment in different contexts are also discussed. In conclusion, it is shown that there is a compelling need for self-assessment research so that its viability for New Zealand can be ascertained.

If a predictor is to be useful in a selection situation it has to be both accurate and acceptable. Chapter Five examines the acceptability of self-assessment as part of the process of conducting the present investigation. The specific aims of the research project are outlined. These are: -

1. An evaluation of the acceptability of self-assessment as a selection tool for New Zealand practitioners.
2. An evaluation of the unique predictive variance offered by self-assessment in applied settings above that provided by selection methods used in New Zealand organisations.
3. An evaluation of possible moderator effects on self-assessment - work criteria relationships.

The results of implementing self-assessment in a sample of settings are described and discussed.

Chapters Six to Nine describe the four settings in which the validity of self-assessment was investigated. Chapter Six reports the study conducted in an electronics factory in which the selection procedure included an interview and a psychological test battery.

Chapter Seven reports two studies which took place in a large food processing and manufacturing organisation involving a sample of seasonal employees. The two year period over which the investigation was conducted involved two consecutive processing seasons. Each season was treated as a separate study mainly due to changes in the structure of the seasonal work.

Chapter Eight reports two studies conducted in a Polytechnic Nursing and Health Studies Department using student nurses. The investigation in this setting was conducted as two studies due to changes in the format of the self-assessment request by the organisation.

Chapter Nine reports two studies conducted in a government department using graduate applicants for career positions. The investigation was conducted as two studies as the selection procedures were changed from a mainly interview based procedure to an assessment centre.

Chapter Ten is a general discussion of the investigation. Conclusions regarding the validity of self-assessment as a predictor in personnel selection in New Zealand in both statistical and practical terms are reported. The influence of individual characteristics on the validity of self-assessment are discussed. Observations regarding the acceptability of self-assessment in New Zealand are made and

suggestions for future research are described.

CHAPTER TWO

CURRENT PERSONNEL SELECTION AND PRACTICE

Current Issues

The effect of legislation on personnel selection research and practice has been profound. Various legal attempts in a number of countries to enforce fair and valid selection practice appear to underlie a reduction in published validity studies, at the same time, encouraging a growing research emphasis on validity generalisation and utility analysis. While current research into selection predictors continues the pursuit for valid and fair procedures, actual selection practice appears somewhat less concerned with these goals.

The single most important influence on modern personnel selection in the United States of America (USA) was the introduction of the 1964 Civil Rights Act and its Title VII (Zedeck & Cascio, 1984). Civil rights legislation in the USA and elsewhere began from concerns over individual rights to privacy and nondiscriminatory treatment (de Wolff & van den Bosch, 1984). The Uniform Guidelines on Employee Selection, produced by the Federal Government in 1978, were aimed at assisting practitioners in the implementation of the legislation (Grant, 1980). Unfortunately, confusing differences existed between the advice outlined in the Federal guidelines and a supposedly complimentary document, the "Principles for the Validation and Use of Personnel Selection Procedures" issued by the Division of Industrial and Organisational Psychology of the American Psychological Association in 1980 (Tenopyr & Oeltjen, 1982). Subsequent litigation has suggested basic principles for some aspects of the validation process (e.g., Hogan & Quigley, 1986; Thompson & Thompson, 1982) although court imposed standards have varied somewhat from case to case (Zedeck & Cascio,

1984). The recent publication of the "Standards for Educational and Psychological Testing" is expected to resolve some of the legal disagreements about the quality of criterion-related validity studies (Klieman & Faley, 1985).

In the United Kingdom the Race Relations Act (1976) and the Sex Discrimination Act (1975) have encouraged occupational psychologists to closely examine personnel practices (Wallis, 1980). Guidelines produced by a working party of psychologists (British Psychological Society, 1978) require an effective job analysis, reliable, valid, and nondiscriminatory selection methods, as well as empirical proof of the relationship between test scores and job performance. It may be, however, that recent legislation has in part reversed some of the effects of these earlier Acts. For example, the Employment Act of 1980 restricts the right to compensation for unfair dismissal, a right that was established in earlier legislation (Shackleton & Anderson, 1987). Whether or not practitioners are complying with existing law regarding personnel selection practices is unclear. Some reports state that organisations are prepared to operate within the law (e.g., Lewis, 1984), while others indicate that employers often are not complying with anti-discrimination legislation (Keenan & Logue, 1985).

In the Netherlands, personnel officers have adopted a code of conduct based on a government commission's recommendations (de Wolff & van den Bosch, 1984). The two major principles of the code are, that selection criteria must be based on the job in question and that selection procedures should be consistent with human dignity. Applicants are entitled to have access to and confidential treatment of personal data and retain the right to lodge a complaint.

Legislation to promote human rights in New Zealand and to establish the Human Rights Commission and the Equal Opportunities Tribunal was enacted in November,

1977. It extended the meaning of unlawfulness as applied to racial discrimination covered by the Race Relations Act 1971. The legislation was intended to promote and monitor equal access to jobs and other goods and services. Essentially, discrimination by reason of sex, marital status, religious or ethical belief, or by colour, race, ethnic, or national origin was made unlawful. Specifically, employers were required to avoid discriminatory selection practices and engage individuals on the basis of job related ability. Unfortunately employers continue to maintain stereotypes on the basis of sex and race (Hesketh, 1984; Hicks, 1984). In addition, the largest number of complaints received by the Commission involve employment and employment related advertising (George & Smith, 1986b).

Although civil rights legislation gave applicants and employees a voice it also has had a number of other noticeable effects. As intended, practitioners have been encouraged to review their procedures in terms of adverse impact as well as validity (Reilly & Chao, 1982). Unfortunately it also appears to have reduced levels of published research. Organisations have either stopped recording selection procedures or avoided publishing validity data in an attempt to avoid litigation (Boehm, 1982). Regulatory pressure also has changed the nature of validity research. For instance, the validation models used by many practitioners and researchers are aimed at serving the needs of judges rather than psychologists (Landy, 1986). In addition, many of the approaches being used to evaluate fairness and bias do not include all possible costs and benefits associated with different types of selection "hits" and "misses" (Dunnette & Borman, 1984).

Regulatory pressure for valid selection methods has encouraged researchers to consider the generalisability of selection predictors using meta-analytic techniques (e.g., Hunter, Schmidt, & Jackson, 1982). The variability in validity coefficients seen

between investigations of similar jobs using identical methods has traditionally been explained by differences in the factor structure of performance in those jobs.

According to this view, as one cannot be sure of the generalisability of a predictor, empirical validation is necessary in each selection situation. The results of meta-analytic studies however tend to disagree with such situational specificity.

Schmidt, Hunter, and colleagues (e.g., Pearlman, Schmidt, & Hunter, 1980; Schmidt & Hunter, 1977; Schmidt, Hunter, & Caplan, 1981) feel that much of the variance in employment test validity is created by statistical artifacts that are unrelated to the true relationship between the test and job performance. When correction for the effect of between-study differences (criterion unreliability, predictor unreliability, range restriction, and sampling error) is made, the remaining validity variance is near zero or zero. This means that if there is sufficient information to ensure that a job is in fact similar to those for which validity distributions are available, an accurate estimate of true validity can be made without conducting a validation study (Pearlman et al., 1980). The situational specificity versus validity generalisation controversy has been discussed and argued at length (e.g., Callender & Osbourn, 1982; Raju & Burke, 1983; Schmidt, Hunter, & Pearlman, 1982). Two very interesting papers, arranged in debate format, provide a useful summary (Sackett, Tenopyr, Schmitt, & Kehoe, 1985; Schmidt, Pearlman, Hunter, & Hirsh, 1985).

Among the many issues investigated by those using validity generalisation techniques is test or predictor fairness. Meta-analytically summarised research shows that single group and differential validity by race is artifactual rather than substantive in nature (Schmidt & Hunter, 1980). When predictors are examined for cultural bias and various statistical artifacts are controlled for, it has been found that differential validity

exists no more often than would be expected on the basis of chance (e.g., Schmidt, Pearlman, & Hunter, 1980).

Group and subgroup differences in ability test scores are mirrored by a corresponding difference in academic achievement and performance on the job (Hunter & Hunter, 1984). Control of adverse impact is a policy decision by the employing organisation. Adverse impact can be dealt with by independent control of the predictor results (Schmidt et al., 1985). The effect of such a correction is minimal as it has been shown that adverse impact can be avoided while maintaining 85%-90% of the productivity gains from valid selection techniques (Schmidt & Hunter, 1981).

Meta-analysis as a method of evaluating predictors and summarising validity coefficients also has provided useful information for the estimation of the costs and benefits of various predictors. One of the major effects on utility computations is the validity coefficient of the predictor or predictors concerned. Utility formulae have been available for a number of years (e.g., Brogden, 1949; Cronbach & Gleser, 1965), but the information required for computation has been difficult to obtain. Recent efforts using rational estimation methods show some promise in avoiding the costly and complicated accounting methods previously required (Cascio, 1987; Hunter & Schmidt, 1982b).

Estimates of productivity gains from using valid selection techniques have produced some impressive results. For instance, yearly productivity gains from improved selection validity in US Federal government jobs is reported to be \$600 million (Schmidt, Hunter, Outerbridge, & Trattner, 1986). Similarly a police department employing 5,000 individuals using cognitive ability tests for officer selection would save \$18 million for each year's hires (Hunter & Schmidt, 1982a). Interestingly,

researchers in this area are also considering the effect of the candidate's decision on utility gains. For example, Murphy (1986) showed that as offer rejection by high scoring candidates necessitates acceptance of lower scoring candidates, estimation of utility gains by present methods may well be an overestimation by 30 to 80 percent.

While the trends discussed above have implications for personnel selection as a whole, the emphasis on improving the validity of individual predictors continues. Recent activity includes the development of new approaches such as "accomplishment inventories" (e.g., Hough, 1984), and further research using more established techniques such as biodata (e.g., Davis, 1984). Comparative reviews evaluating the validity, utility, and practicality of various approaches (e.g., Boehm, 1982; Hunter & Hunter, 1984; Monahan & Muchinsky, 1983; Reilly & Chao, 1982) carry on the tradition of providing information on relative effectiveness. Although selection procedures from the candidate's perspective are occasionally considered (e.g., Reilly & Chao, 1982; Robertson & Makin, 1986) most of the research examines selection from the employers' point of view.

Predictor Research and Practice

Current research continues the convention of reporting validity information on a wide range of techniques. The following section provides an illustration of predictor research by describing and discussing some examples of recent investigations.

Psychological Tests

Psychological tests are used extensively in personnel selection, both in New Zealand (Hesketh, 1974) and other countries (Grant, 1980) although legislation may have reduced their use. For example, according to a 1985 survey of 7000 USA human

resources managers, around 17% of companies use psychological tests for selection (Aberth, 1986). The survey showed that fewer organisations were prepared to risk evaluation by the Equal Employment Opportunity Commission. In comparison, test usage in Britain has not been limited in the same way although testing has not played as large a part in selection as in the USA (Anderson & Shackleton, 1986).

The thousands of available tests can be categorised into traditional administrative categories (e.g., speed, power, group, individual, etc.), the type of behaviour sampled by the test (Cleary, Humphreys, Kendrick, & Wesman, 1975), or the test content. Ghiselli's (1966) summary of thousands of studies over a sixty year period classified tests into five main ability categories.

- a) Tests of intellectual abilities such as those measuring intelligence, immediate memory, and arithmetical abilities.
- b) Tests of spatial and mechanical abilities.
- c) Tests of perceptual accuracy such as those involving name and number comparison, cancellation, and pursuit activity.
- d) Tests of motor abilities using tracing, tapping, and dotting tasks as well as testing finger, hand, and arm dexterity.
- e) Tests of personality traits, a category which includes interest inventories.

Hunter and colleagues (e.g., Hunter & Hunter, 1984) combined intellectual, spatial, and mechanical abilities into a single category, including tests of inductive and deductive reasoning. A reanalysis of Ghiselli's (1973) data using this categorisation found that as the complexity of the content of jobs decreased so did the size of the validity coefficients (Hunter & Hunter, 1984). These authors showed that an optimal combination of cognitive ability and psychomotor tests produced validity coefficients

ranging from .62 (salesperson) to .28 (sales clerk), the latter being the only coefficient falling below .43. They concluded that the mean validity of an optimal composite of ability measures was .53 for entry level jobs. The same mean validity was only exceeded by work sample tests (.54) in those situations where tests were used for certification or promotion. General cognitive ability has been described as a predictor of a range of jobs including so called "manual" and "mental" jobs (Hunter, 1986). A validity generalisation system is being developed to predict success in virtually all occupations using tests such as the General Aptitude Test Battery (Hawk, 1986).

A range of personality tests are used in selection (Monahan & Muchinsky, 1983). Examples include the Sixteen Personality Factor Questionnaire (16PF; Cattell, Eber, & Tatsuoka, 1980), the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway, & McKinley, 1967), the California Psychological Inventory (CPI; Mergargee, 1972), the Edwards Personal Preference Schedule (EPPS; Edwards, 1957), and the Guilford-Zimmerman Temperament Survey (Guilford & Zimmerman, 1956). The 16PF seems to be the most popular personality test among British occupational psychologists (Tyler & Miller, 1986) and is widely used in New Zealand (Hesketh, 1974) and Australia (Smith, Dowling, & Barry, 1985). An evaluation of individual assessment practices in the USA showed the 16PF to be the most commonly used followed by the Guilford-Zimmerman Temperament Survey, the CPI, the MMPI, the Myers Briggs Type Indicator, and the Edwards Personal Preference Test (Ryan & Sackett, 1987). Over 54% of police departments in the United States use personality tests, the most popular of which is the MMPI (Parisher, Rios, & Reilly, 1979).

In general, personality tests are poor selection devices with validity coefficients of around .15 (Schmitt, Gooding, Noe, & Kirsch, 1984). Although job analyses show

affective and interpersonal characteristics are important for some jobs, personality tests have generally failed as useful predictors of performance (Thornton & Byham, 1982). Guion and Gottier (1965) concluded from an extensive review, "there is no generalisable evidence that personality measures can be recommended as good or practical tools for employee selection" (p. 159). Personality tests are open to biased responding in that test takers can, and will, answer in terms of what they think an organisation wants. Personality profiles describe what sort of personality an applicant feels is appropriate to the occasion rather than describing the personality of the applicant. Applicants are easily able selectively to distort personality "scores" without such distortion registering on so called "lie scales" (Elliot, 1981).

Recent reports of frequency of usage and predictive validity for projective personality tests are noticeably lacking. One exception was a survey of United States individual assessment practices which showed half the respondents used projective tests in personnel selection (Ryan & Sackett, 1987). The survey identified the most popular instrument as a sentence completion task, either self-developed or commercially available, such as the Miner Sentence Completion Scale (MSCS; Miner, 1977). Other tests mentioned were the Rorschach (Taylor & Nevis, 1957) and the Thematic Apperception Test (TAT; Roe, 1953).

Although research occasionally reports significant predictive validity (e.g., Grant, Katkovsky, & Bray, 1967), projective tests as predictors have little practical value (Kinslinger, 1966). A recent review reported a mean validity of .28 based on six studies and concluded that projective tests lack the basic requirements of sufficient reliability and validity to be a viable alternative to other selection methods (Reilly & Chao, 1982). The MSCS was the only test that had demonstrated usable validity in a selection setting.

Interviews

The selection interview is probably the most widely used personnel selection technique (Guion, 1976). Since the first comprehensive review of the interview (Wagner, 1949) the discouraging conclusions reported have become persistent in research and later reviews of this area (e.g., Mayfield, 1964; Schmitt, 1976; Ulrich & Trumbo, 1965; Wright, 1969) although some positive findings do exist (e.g., Arvey, Miller, Gould, & Burch, 1987).

Reviewing interview research from 1975, Arvey and Campion (1982) found that while research continues to be microanalytic in approach, methodologies are becoming more sophisticated and realistic. Their review showed investigations of possible bias in the interview have increased and a broader range of variables are being considered (e.g., nonverbal behaviour, interviewee perceptions, and training). Other reports point to more emphasis on the process of the interview, especially on the content and the way information is used by interviewers (e.g., Dougherty, Ebert, & Callender, 1986; Kinicki & Lockwood, 1985; Raza & Carpenter, 1987; Zedeck, Tziner, & Middlestadt, 1983). There is also a growing trend towards applying theories of social cognition to the interviewing process (Guion, 1987).

Webster (1982) describes the interview as unique in that it not only collects information, but is aimed at making a hiring decision which in itself exerts considerable pressure on the interviewer. He feels that research which ignores what he terms "decisional stress" and how it interferes with rationality, cannot adequately describe the selection interview. The reality of an interview situation may involve selective impression formation, implicit personality judgments, a pressure to hire, and

possible negative consequences for the interviewer. Webster feels that these factors are all too often ignored in interview research.

A relatively recent development has been the situational interview (Latham, Saari, Pursell, & Campion, 1980). This approach involves asking "what would you do if?" questions, which are based on critical incidents of job behaviour. Candidates' answers are rated by the interviewer. Research has shown such interviews to be valid predictors of future performance (Latham & Saari, 1984). A specific example using salespersons reported a validity coefficient of .45 against a criterion of sales productivity (Weekley & Gier, 1987). Situational interviews seem to fall somewhere between the traditional selection interview and an actual work sample. While not actually performing a task, candidates are required to express verbally responses to problems which are important components of job performance.

Of particular interest is a growing body of research on how the interview is seen from the candidate's perspective. A candidate has the opportunity to form overall impressions and gather other information during selection interviews. Such information has practical importance when job offers are considered. Schmitt and Coyle (1976) showed that a candidate's overall impression of an interviewer is an important factor in an employment decision. In a study of seasonal worker applicants, how competent, personable, or informative interviewers appeared, strongly influenced attitudes toward the organisation (Liden & Parsons, 1986).

How likeable an interviewer appears also seems to positively affect the willingness of graduates to accept a job offer (Alderfer & McCord, 1970; Keenan, 1978). Evidence that interviews can negatively affect intentions to accept jobs has been reported by Herriot and Rothwell (1981). They suggest that a negative reaction could be due to

interviews not meeting candidates' expectations. Other work shows that applicants tend to react more favourably to interview content concerning general qualifications for the job (Taylor & Sniezek, 1984). Training to improve interviewee performance also has been shown to be useful (Mathews & Fawcett, 1984). The role of the interviewee is apparently not a passive one and deserves further attention.

Application Form Information

Completion of an application form is a requirement faced by most applicants. The details provided can be used for preliminary screening, to provide background information for an interview, or weighted and combined to predict some form of job success. The extensive research literature suggests that the weighted application form is one of the best predictors of tenure (Muchinsky & Tuttle, 1979) and in many instances, job performance (Tenopyr & Oeltjen, 1982). Reilly and Chao (1982) reported a range of cross-validated validity coefficients for five types of criteria (tenure, training, ratings, productivity, and salary) of between .32 and .46 with an average of .35. Weighted application forms do have their limitations however. They tend to be specific to situations, suffer decay in validity over time, and require a large sample (400-1000) for development (Hunter & Hunter, 1984).

The methods used to weight application form information are many and varied (Freeburg, 1967). Earlier differential weighting approaches such as the horizontal percent method (England, 1971) and correlational analysis, weight items individually without considering other items on the form. Weighting methods which explore between item and item-criterion relationships typically involve multivariate analysis such as linear discriminant function analysis (Klecka, 1984). Research has shown linear discriminant analysis to produce comparable results to horizontal percent weightings with the former approach providing more comprehensive statistical

information (Smith & George, 1987). Another approach, rare response scoring, assigns weights based on item frequency but has shown little stability and validity (Telenson, Alexander, & Barrett, 1983). An investigation using clustering techniques to generate homogeneous "keys" for analysis based on multiple correlations reported encouraging results (Matteson, 1978).

Asher (1972) categorised application form data into two types, "hard" and "soft". Hard data are historical and verifiable information, soft data are information that is private and difficult to verify. The Biographical Questionnaire (BQ; Owens & Schoenfeldt, 1979), is a well researched instrument using mainly soft data. The BQ has a stable factor structure (Eberhardt & Muchinsky, 1982a) and useful predictive validity (Davis, 1984; Neiner & Owens, 1985).

Although applicants seem to report soft data less accurately than hard data (Shaffer, Saunders, & Owens, 1986), they do allow an examination of the conceptual basis of factors underlying past behaviour (Owens, 1971) and the possible effects of past experiences (Eberhardt & Muchinsky, 1982b). Focusing solely on empirically derived relationships between hard data and criteria seems less acceptable in the light of this recent work using the BQ despite concerns over the accuracy of responses to soft items.

Weighted application form information appears to be used exclusively by employers though not by a large proportion of them (Makin & Robertson, 1986). Candidates rarely have access to details of how various application form information is related to work criteria. An example of how this information can be used was demonstrated by a New Zealand Polytechnic which told prospective candidates of factors that were related to course dropout rates (George, 1984). For instance, having responsibility for

children was associated with dropping out of the course. The Polytechnic also used these data to plan support services to compensate for certain effects identified by the same information.

References, Letters of Recommendation

Personal history is not assessed solely by self-report as in application forms. Descriptions of applicant background can come from buddy or peer ratings and references or letters of recommendation. Muchinsky's (1979) review of the limited research available concluded that while validity is generally poor, properly constructed reference reports can be a source of incremental validity. This conclusion is supported by a study showing that a carefully developed reference form used by head teachers is an effective predictor of navy officers' training performance (Jones & Harrison, 1982). Muchinsky (1979) also felt that references can provide useful, additional information when screening applicants although legal pressures may reduce their use overall.

Knouse (1983) showed that references giving specific information (e.g., examples and numerical data) have a positive effect on candidate ratings and that references containing specific examples of a candidate's behaviour tend to enhance a reference writer's credibility. He also found mixed results for the presence of a negative statement in a reference. An unfavourable statement increases an applicant's employability but decreases estimates of overall ability. These trends do not vary greatly for different types of referees, different lengths of time, and different conditions of acquaintance (Baxter, Brock, Hill, & Rozelle, 1981).

Amendments to United States law allowing candidates the opportunity to waive their right to inspect references have been investigated. References which are confidential,

as opposed to public, tend to be written in a less favourable way (Ceci & Peters, 1984) and are treated more favourably by individuals using them for selection (Shaffer & Tomarelli, 1981).

Work Sample Tests

A work sample is a test based on job skills which are crucial to job performance. The high level of similarity between predictor and criterion is based on the principle of behavioural consistency (Wernimont & Campbell, 1968). A work sample may measure past learning and experience or predict potential to learn in the future, as in the case of trainability tests (Robertson & Downs, 1979). Current research on work samples has shown that they are among the best predictors of job performance, especially in the case of entry level positions (Hunter & Hunter, 1984). Their predictive power when compared to traditional tests is amply illustrated in a growing body of research (e.g., Gordon & Kleiman, 1976; Mount, Robertson & Kandola, 1982).

There are many different types of work sample. Asher and Sciarrino (1974) distinguish between verbal work samples (e.g., in-basket tests or leaderless group discussions) and motor work samples which require physical manipulation. Robertson and Kandola (1982) divide work sample content into psychomotor, individual - situational decision making, job related information (e.g., pencil/paper tests), and group exercises. Work samples are expensive to develop and maintain as they are job specific and may need regular modification to reflect changes in job content. Work samples have been described as fair and able to provide realistic job information (Farr, O'Leary, & Bartlett, 1973) which positively affects self-selection trends (Downs, Farr, & Colbeck, 1978).

Assessment Centres

Assessment centres use a number of selection techniques in which work sample tests feature prominently. Apart from using multiple methods and assessors, a distinctive feature of assessment centres is the subjective consideration and judgmental combination of information. Exact weights are not used but the various types of information are given more or less importance depending on assessors' estimates of relevance to the job in question (Thornton & Byham, 1982).

Assessment centres were first used in military settings to select officers for the German and British armies and covert operations staff for the American Office for Strategic Services (Wiggins, 1977). Initial non-military uses included the selection of Australian executive trainees (Taft, 1948), the British Civil Service Selection Board (Anstey, 1971, 1977), the American Telegraph and Telephone Company's Management Progress Study (Bray & Grant, 1966), and the US Veterans Administration programme to support the training of clinical psychologists (Kelly & Fiske, 1951). Apart from managerial selection, assessment centres have been used as training and selection criteria to assess professional and academic competence, for selection of non-supervisory personnel such as salespersons, engineers, scientists, and police recruits, and as a basis for career evaluation (Thornton & Byham, 1982).

A vast amount of research has been conducted and reported since these early times reflecting a range of applications in industrial and other settings (e.g., Byham & Wettengel, 1974; Gaugler, Rosenthal, Thornton, & Bentson, 1987). Of particular note is the debate over the content and construct validity of assessment centres.

Generally an assessment centre is developed on the basis of an extensive analysis of job content (Thornton & Byham, 1982). The validity of the centre depends on the content analysis represented by behavioural constructs or dimensions of performance

considered important for the target position. Some researchers have questioned the validity of these dimensions. In factor analytic terms, assessment centre ratings do not produce factors corresponding to the dimensions rated but rather factors related to the exercises used (Bycio, Alvares, & Hahn, 1987; Sackett & Dreher, 1982). While this does not destroy the job-related nature of the procedure (Neidig & Neidig, 1984) it does highlight the lack of evidence showing why assessment centres actually work as predictors (Klimoski & Brickner, 1987; Zedeck, 1986).

Assessor judgment processes do not appear to coincide with assessment centre architects' assumptions regarding the role of dimensions or how overall ratings are arrived at (Russell, 1985). Sackett and Dreher (1984) suggest eliminating global dimensions and focusing instead on the identification of managerial roles and the design of exercises to simulate role performance. Each exercise is treated as a work sample of managerial behaviour and scored as a separate predictor (Robertson, Gratton, & Sharpley, 1987). Assessment centres are treated as a collection of predictors, not processes producing a single overall score. Another suggestion involves designing each exercise to yield behaviour that will allow the rating of a large number of abilities (Bycio, Alvares, & Hahn, 1987). Instead of focusing on a particular role, every exercise will attempt to measure virtually all managerial skills identified in the target job.

Assessment centres are shown to be valid predictors of work performance, similar to cognitive tests and work samples, with meta-analytically derived validity coefficients of around .40 (Schmitt et al., 1984). Assessment centre based predictions appear equally valid for male and female managers (Ritchie & Moses, 1983). A recent review reported a mean validity coefficient of .37 once corrected for sampling error, restriction of range, and criterion unreliability (Gaugler et al., 1987). Significant

moderator effects detected in this review showed that "validities were higher when the percentage of female assessees were high, when several evaluation devices were used, when assessors were psychologists rather than managers, when peer evaluation was used, and when the study was methodologically sound" (p. 493).

Clearly, not all comments on the assessment centre have been positive. There are concerns over sloppy development as more organisations use this approach (Dunnette & Borman, 1979). Klimoski and Strickland (1977) did suggest assessment centre evaluations only predict management progress and do not assess managerial performance. This view is supported by recent research which highlights a number of possible distortions in criteria used to evaluate assessment centre predictions (Turnage & Muchinsky, 1984). The reliability of assessor ratings are often poor (e.g., Hinrichs & Haanpera, 1976; Jones, 1981) and assessment centres may have difficulty predicting some aspects of the leadership role, such as task orientation (Cunningham & Olshfski, 1985). Criticisms of the process involved in assessment centres have led to calls for a careful evaluation using, for example, information processing, categorisation and social cognition, and group dynamic approaches (Zedeck, 1986).

Other Predictors

In addition to those discussed above there are a number of other predictors and related issues currently receiving attention in the research literature. An example is computerised tailored testing, sometimes known as adaptive testing. This approach is based on the idea that different sets of items with different properties may be best for different people. Pencil and paper tests have been criticised for their excessive administration time, poor differentiation among people of extreme ability, limited capacity for measuring some types of ability, clumsy and error-prone scoring, cost, and vulnerability to theft and compromise (Hakel, 1985). Computerised testing can

incorporate tailored testing where testee responses have an effect on subsequent items presented. This allows a more precise measure of ability, significantly fewer items, and therefore less testing time relative to conventional paper and pencil tests (Tenopyr & Oeltjen, 1982).

A further development of testing described by Guion and Ironson (1983) is tailored measurement. Given the criticisms made of conventional tests and the difficulties involved in developing truly parallel forms for adaptive tests, Guion (1987) suggests a test could be derived for each situation:

"Given a pool of calibrated items, one could develop a series of tests with similar information curves. These could be very nearly as short as the tailored tests, could be changed often to minimise threats to test security, and could provide interchangeable scores in the form of estimates of the underlying trait" (p. 204).

Computerised tests can be designed to be convenient, accurate in scoring, and secure although there is potential for abuse. With increasing interest in computer based testing, item response theory may play a greater role in personnel selection testing, especially in relation to sequential testing, test development, and nonparallel test comparison (Hakel, 1985).

Graphology, or handwriting analysis, is routinely used in personnel selection by an estimated 85% of companies in Europe and around 3000 firms in America (Rafaeli & Klimoski, 1983). These amazingly high estimates of usage are not only surprising but also fly in the face of research evidence. Handwriting analysis has been referred to as brain-writing (Frederick, 1965) with samples appearing to offer a rich variety of individually produced information that is relatively stable over time (Fluckiger, Tripp, & Weinberg, 1961). Recent investigations of validity demonstrate little evidence of practical value and few, if any, significant relationships (Rafaeli & Klimoski, 1983). Graphologists' judgments based on spontaneously produced text show small positive

validity coefficients as do non-graphologists using the same information (Ben-Shakhar, Bar-Hillel, Bilu, Ben-Abba, & Flug, 1986).

A number of reports have appeared dealing with methods of predicting specific aspects of different jobs. Two examples involve the slightly futuristic concept of genetic screening and the ever present concern with deception. To screen out individuals who are hypersusceptible to toxins in work environments it seems that genetic screening devices may soon become part of employment screening procedures (Olian, 1984). Olian's paper showed that while the validity of such devices can be evaluated as for other tests, the legal status of such instruments is presently ambiguous and may require specific policy.

Devices designed to detect deception are being widely used in American corporations (Belt & Holden, 1978) and government agencies (Working Group of the Scientific Affairs Board, 1986). A review and analysis by Sackett and Decker (1979) concluded that estimates of polygraph accuracy taken from criminal investigation contexts cannot be generalised to the employment setting given the considerable chance of innocents being judged guilty. For example, voice analysis instruments may only have value in a fear inducing role as they are unable to detect deception at a greater than chance level. Similar conclusions exist for paper and pencil predictors (Sackett & Harris, 1984). As a source of selection data it appears detectors of deception are of little predictive use, although like other methods, such as personality tests, the attraction may lie more in the mere use of these devices rather than the quality of the information they produce (Munchus, 1986).

The effectiveness of individual predictors continues to be a major concern of researchers and practitioners. Comparative reviews not only consider the relative

validity of different approaches but also evaluate the predictability of different criteria (e.g., Gaugler et al., 1987; Hunter & Hunter, 1984; Schmitt et al., 1984). These results show that many predictors display different validity coefficients for different criteria. The following section discusses the characteristics of the various criteria commonly used in predictor research.

Measures of Work Performance

As illustrated by Smith's (1976) three-dimensional framework for classifying performance measures, the range of criteria against which selection methods are compared varies widely. The measures used in most cases tend to be based on judgmental ratings although promotional indices, turnover, and absenteeism records are also used as measures of individual and organisational effectiveness. It is argued that a single measure of work performance will not necessarily reflect the multi-dimensionality of work performance (Guion, 1961; Smith, 1976); so where possible a range of measures should be used. While such an approach must be tempered by the sort of data that are available in applied settings, a range of possible criterion measures is often available.

The most widely used indicators of performance are judgmental and usually based on graphic rating scales (Landy & Farr, 1983). The problems of performance appraisal are well documented (e.g., Drenth, 1984; Landy & Farr, 1983; Smith, 1976), yet the amount of research using such criteria is considerable, based on the sometimes questionable assumption that such a criterion itself is fair and valid (Linn, 1984).

Promotion as a criterion of work performance has been used in a number of studies despite the possibility that promotion decisions may be affected by factors other than

performance. Political expediency, organisational structure, labour market conditions, word of mouth, and other informal evaluations (Campbell, Dunnette, Lawler, & Weick, 1970; Smith, 1976) can contribute to promotional decisions. The use of promotional progress as a criterion had been shown to yield significantly higher validity coefficients than when performance appraisals were used for a range of predictors. This suggests that promotional progress may be a more reliable and valid measure of job performance (Meyer, 1987).

Meta-analytic studies have shown that criteria such as status change and earnings produce higher validity coefficients on average than performance ratings (Schmitt et al., 1984). Unfortunately there appears little consistency in how promotional criteria are actually used. As an example, one meta-analytic review combined change in salary over time, absolute salary level, number of promotions, absolute job level, and turnover as one category of criteria titled "Career Advancement" (Gaugler et al., 1987).

The concept of absenteeism as a measure of efficiency has been described as "seductive". "Who could argue with a straight face that an absent employee was a productive one?" (Landy & Farr, 1983, p. 28). From the organisational point of view, absenteeism is regarded as a cost to efficiency with massive savings possible if it could be reduced (e.g., Mowday, Porter, & Steers, 1982; Steers & Rhodes, 1978). Absenteeism has been considered from an individual perspective as withdrawal in reaction to stress (Hill & Trist, 1953) and as a reflection of behaviour at the group level where it is treated as a social phenomenon (Chadwick-Jones, Nicholson, & Brown, 1982).

Despite some debate over how absenteeism is investigated, there seems to be general agreement that it is an important indicator of efficiency (e.g., Chadwick-Jones et al., 1982; Hammer & Landau, 1981; Keller, 1983; Landy & Farr, 1983). Absenteeism has been measured in a variety of ways but the conditions of applied research often require that measures need to be empirically derived (Huse & Taylor, 1962). Duration measures, such as the number of days workers are absent, is often used although a common alternative measure is a frequency index which is based on the number of absences in a given period rather than the amount of days or shifts missed (Folger & Belew, 1985). The frequency measures tend to exhibit higher reliability than other measures (Muchinsky, 1977). Unfortunately work comparing absenteeism measures is sparse and there tends to be no uniformly accepted classification scheme (Mowday et al., 1982).

Turnover is often a costly and disrupting influence for an employing organisation, encouraging a large number of research studies (Muchinsky & Morrow, 1980; Steers & Mowday, 1981). Turnover has been shown to be systematically related to job performance (Wells & Muchinsky, 1985) and has been often used in organisational research to represent performance (Landy & Farr, 1983).

The definition of turnover has proved problematic. Various approaches include using a dichotomy of "stayers" versus leavers (e.g., Farris, 1971) or focusing on the quality and how replaceable the leaving employee was (Dalton, Krackhardt, & Porter, 1981). Dalton, Todor, and Krackhardt (1982) have suggested separating turnover frequency (i.e., the number of separations) and turnover functionality (i.e., the nature of separations), so dividing employee samples into stayers, voluntary quits, and dismissals (e.g., Parsons, Herold, & Leatherwood, 1985).

The distinctions researchers may feel are appropriate are not always clear in practice however (McEvoy & Cascio, 1985). Much turnover research relies on information in personnel files and important differences have been found between the official turnover reasons noted on file and those given by past employees approximately six months after termination (Lefkowitz & Katz, 1969). How does one classify cases where individuals are "encouraged" by the organisation to resign or other individuals who take early retirement? One option, defining turnover as a permanent movement beyond the boundary of an organisation (Macy & Mirvis, 1983), may minimise the effect of shades of interpretation but would omit other important information.

As an additional indicator of work performance, the difference between the meaning attached to dismissal for work related reasons and resignation for reasons outside the job is clear. The former reflects unacceptable performance. The latter may not have anything at all to do with work. The functionality of turnover is more critical for organisation effectiveness than is turnover frequency (Hollenbeck & Williams, 1986). Dividing subject samples into those who stay and those who go would appear to neglect information related to job performance. Studies that have separated employees leaving into voluntary and non-voluntary sub-groups have found significant differences in predictability and job satisfaction (Muchinsky & Tuttle, 1979). On balance, it would appear that turnover should be measured in terms of functionality. Subjects can be classified as either "Stayers", "Voluntary Turnovers" (e.g., resignation, retirement), and "Non-voluntary Turnovers" (e.g., dismissal).

Personnel Selection in New Zealand

Information on personnel selection practices and research in New Zealand is fragmented and sparse (Inkson, 1987). No definitive source, such as a comprehensive

survey, appears to be available with information often unpublished (e.g., Toh Eng Lim, 1981) or embedded in reports which focus on more general or peripheral issues (New Zealand Institute of Personnel Management, 1979). A description of the New Zealand situation can therefore only be regarded as subjective.

In general terms personnel selection in New Zealand seems to involve unstructured interviews supported by an application form and references or referee comments although various other techniques are sometimes used for particular jobs. An examination of staff selection practices in 40 Canterbury businesses indicated that the interview, application form, and reference were the common sources of information (Henderson, 1987). This study showed that some organisations also used medical examinations (n = 6, 15%), personality tests (n = 2, 5%), a work sample (n = 1, 2.5%), and graphology (n = 1, 2.5%). Over 20% of these Canterbury companies, on occasion, used other techniques such as cognitive ability tests, clerical tests, and work sample tests. Interestingly, none of the 40 respondents in Henderson's survey had investigated or were intending to investigate the validity of their selection methods. The survey also showed that the use of job analysis techniques appear restricted to the training field with few practitioners using other than a brief job description as a basis for selection.

A survey aimed at describing how managers were selected in New Zealand showed that 83% of the 63 responding organisations regarded the interview as fairly or very important to managerial selection (Toh Eng Lim, 1981). Unfortunately the range of alternative selection methods offered in this survey was restricted to psychological tests, external consultants, and structured and unstructured interviews. Other techniques reportedly used by respondents were peer nomination, "head hunting", and

evaluation of qualifications. The same study revealed that only one out of 63 organisations had used an assessment centre for selection.

Assessment centres in New Zealand are used for officer selection in the defence forces, selection of diplomatic and administrative trainees for the Department of Foreign Affairs (Shouksmith & Petersen, 1982), and for selection of career officers in the Department of Trade and Industry (George & Smith, 1986a). There may be other applications of this approach in private industry but reports of such instances do not appear to be publicly available or identified in survey investigations.

Apart from the common interview based approach, there are a number of reported exceptions worthy of mention. As for other jobs, selection of trade apprentices in New Zealand is still largely based on an informal interview (Dakin & Inkson, 1985). One particular organisation, NZ Steel, has reported apprentice selection procedures which include an evaluation of background information, interviews, and five aptitude tests (Jenkins, 1985). While this approach may not appear remarkable, the fact that these procedures were based on job and task specifications, and validated against technical institute exam results and on-job performance ratings is unusual in the New Zealand setting (e.g., Henderson, 1987). In addition to selecting high performing apprentices, NZ Steel have had only one individual out of 85 prematurely leave the training programme in four years. The level of attrition expected, based on labour turnover in the company, was 25 (Jenkins, 1985).

The New Zealand Armed Forces use a variety of techniques to select service personnel most of which are based on extensive job analyses (Toulson & Williams, 1979). Little information, especially psychometric details, are available on the exact nature of the selection methods. Toulson and his colleagues (Phipps & Toulson, 1982; Toulson, 1983;

Toulson & Williams, 1979) show that selection is based largely on the experiences of the British and American armed services. Requirements of the various branches vary from the extensive training and evaluation techniques to select combat air crew, to the aptitude tests used to classify large numbers of army trade personnel. Officer selection is based on a British selection board model (assessment centre) and involves a variety of tests (cognitive ability, aptitude tests, leaderless and leader designated group exercises, physical tests, problem solving exercises, etc.), and panel interviews. Cadets and other ranks are selected using ability tests, interviews, and in some cases interviews by psychologists (W. McEwan, personal communication, May, 1987).

A rather unique line of New Zealand selection research involves expedition members for the International Biomedical Expedition to the Antarctic. One particular study evaluated biographical, clinical, and psychometric predictors (personality, cognitive ability, and a video interpersonal distance measure). Criterion measures included self, peer, and observer adaption measures made during field work (Taylor & McCormick, 1985). An analysis of predictor - criterion relationships produced 96 correlations, 19 of which were significant at the .05 level. Such a large number of correlations based on a sample size of 12 makes it difficult to draw firm conclusions. The authors did state, however, that "the best predictors of Antarctic performance were polar experience, age, clinical ratings that included subject interviews and the HSCL (Degrogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974), and [the] SACL (MacKay, Cox, Burrows, & Lazzerini, 1978) stress scales" (p. 13).

A disturbing characteristic of New Zealand personnel selection is the use of personality tests for selection by employers and management consultants (Bull, 1974; Dakin & Inkson, 1985; Gimpl & Dakin, 1984). No recent data are available, but an earlier survey by the New Zealand Institute of Personnel Management (Hesketh, 1974)

showed the 16PF and Eysenck's Personality Inventory (Eysenck & Eysenck, 1973) were the most typical. A recent article damning personality tests as selection instruments (Smith & George, 1986a) received a substantial amount of criticism from practitioners who felt tests such as the 16PF had a valuable role to play in selection procedures (e.g., Townsley, 1986). Interestingly, no critics were able to supply evidence of validity for the tests they were defending.

Practitioners and researchers in New Zealand seem to have quite different views of what selection methods are valid. Dakin and Armstrong (1986) found that New Zealand personnel consultants' beliefs as to the validity of a range of selection techniques were statistically unrelated to validities reported in the literature. The consultants involved were most likely to choose an interview and an application form. Their least likely choices were a cognitive ability test and a work sample. New Zealanders seem to have a deep mistrust of tests of ability and aptitude (McLennan, Inkson, Dakin, Dewe, & Elkin, 1987). This is not surprising when so few organisations deal with psychologists. A 1972 investigation found that only 7% of personnel managers in New Zealand organisations would consider employing a psychologist compared with 44% in Australian companies and 60% in the US (Hines, 1972). A 1983 study found little change in these earlier results. Only 16% of companies used psychologists, mainly as consultants (Bull & Spicer, 1983). It is little wonder that practice does not reflect research when those versed in the literature rarely have the opportunity to practice.

While the dearth of personnel selection research in New Zealand is evident, the lack of attention to selection from the candidate's perspective in this and other countries is also apparent. Given that applicants have had the opportunity to become quite knowledgeable about their own abilities and attributes under a variety of conditions, it

would seem a source of potentially valuable information is being ignored by many of the more traditional predictors. An understanding of the problems and possibilities of self-assessment is best approached from an appreciation of the psychology of the self as a basis from which the process of self-assessment can be considered. Accordingly, Chapter Three looks at the self and self-assessment from a psychological perspective.

CHAPTER THREE

THE PSYCHOLOGY OF SELF

The curiosity individuals have in their own thoughts and behaviour is well represented in psychology in its tradition as a scientific profession (Vallacher, 1980). The person's sense of self is one of the most intensely studied issues throughout the history of empirical psychology (McGuire, 1984) producing a wide diversity of perspectives and an enormous volume of research. This chapter does not attempt to provide a comprehensive review of the area but is intended to portray a psychological perspective of the self and issues related to self-assessment. The background to current research on the self will be described and recent trends discussed. Our ability to know and assess ourselves will be considered and the factors affecting self-assessment validity explored.

According to Greenwald (1981), the origin of current psychological interest in the self is illustrated in the work of four individuals representing widely differing backgrounds and schools of thought. Freud talked of the incessant stream of self-reference, Clararede highlighted the important role of the self in memory, Bartlett introduced the concept of schema, and Koffka described the self as a central sub-system of the ego. They all believed that ordinary voluntary memory depended on a high level of organisation. With the exception of Bartlett who avoided accepting what he termed an "intangible and hypothetical self", these psychologists believed that the organisation of memory was largely due to the self (Greenwald, 1981).

This early emphasis on the organisation or durable structure of the self is still prevalent in current theoretical and empirical investigations. Network models of

memory suggest that information about the self is stored in the form of propositions (Bower & Gilligan, 1979), as representations of the meaning of an event rather than the exact structure (Anderson, 1985). Propositional networks can be portrayed as sets of nodes connected by pathways that represent predicate relationships. This view regards the self-concept as a part of an organised system of concepts concerning an individual's social and physical world (Kihlstrom & Cantor, 1984). The self can also be viewed as a list of items or features derived from a lifetime of experience (Rogers, Kuiper, & Kirker, 1977). This approach depicts the self as a hierarchical category structure, the elements of which are traits, values, and memories of specific behaviours (e.g., Rogers, 1981).

Hoelter (1985) proposes a structural model that considers identities as positional elements that can be located within a dimensional space defined by personal attributes. Identities are self definitions or roles arising from group membership as part of the social networks within which people interact. Personal attributes are dimensions of meaning for the identities one holds.

The information processing perspective states that individuals encode and represent information using internal cognitive structures called schemata (Markus, 1977). Schemata or knowledge structures are conceptually related frameworks of information gained from experience that guide and are continually updated by the processing of information (Markus, 1983). Self-schemata are developed by individuals to explain and understand their social experiences (Markus & Sentis, 1982). They are derived from specific events and situations involving the individual as well as the information gained from repeated categorisation and subsequent evaluation by the self and others. The self can be regarded as an attitudinal schema as it is not only an attitude object but is actively involved in acquiring and retrieving knowledge (Greenwald & Pratkanis,

1984). Self-schema exist in dynamic interdependence with the environment, performing the dual role of mediating and interpreting behaviour (Markus 1980). Their influence is shown when individuals resist information that does not agree with prevailing self-schemata (Markus, 1983).

The centrality of the structural features of the self-concept is one of the major aspects of self research that is slowly changing. Gergen's (1984) excellent discussion of evolutionary trends in this area focuses on three aspects. Firstly, there is the move from structure to process where the concern with perpetual enduring entities, such as found in trait research, is declining. Instead there is a focus on interactionism as in the self-presentation literature. Secondly, the mechanistic view of the self as a product of environmental influence in the behavioural tradition is moving to a description of the self as an agent. Bandura's (1978) notion of reciprocal determinism is a prominent example which deals with the concept of agency. Thirdly, there is a change in the traditional belief in individually centred self-knowledge where individuals could tap the repositories of their minds for information. This contrasts with the modern emphasis on placing psychological knowledge in the social sphere. What had passed as knowledge of the self is regarded as a product of social interchange as emphasised by social comparison theory (Festinger, 1954).

Structure to Process: Strategic Self-Presentation

The traditional idea of self-knowledge is intimately connected with the concept of static structure, something that persists over time. Explanations of constructive processes do not fit easily with durable entities or organisations. The search for stable elements of the self-concept that lie beneath the continuity of behaviour are seen in trait research and behavioural approaches. Gergen's (1984) analysis of the

movement from a structural approach to a focus on process suggests measures of individual differences or traits do not reflect the qualities of the individual but represent the accumulated effects of past experience. Personality measures are just shorthand indexes of exposure to past situations and as such render the idea of personality traits obsolete. All human action can be considered a partial result of situational influences. Individuals and situations interact in a reciprocal sense. One area which illustrates this move to treating the self as part of an interactive process is strategic self-presentation.

Strategic self-presentation tactics are "those features of behaviour affected by power augmentation motives designed to elicit or shape others attributions of the actor's dispositions" (Jones & Pittman, 1982, p. 233). Goffman (1969) describes such tactics as a "performance" incorporating "all the activity of a given participant on a given occasion which serves to influence in any way any of the other participants" (p. 14). Presentation strategies are often a reaction to others who have social power or control desired resources, the aim being to increase personal power that will allow an immediate, or even future action, to result in a favourable outcome (Arkin, 1980). Such strategies are not necessarily false or distorted but often involve selective disclosure and emphasis rather than deceit and simulation. Jones and Pittman (1982) offer a taxonomy of presentation strategies as a framework for further investigation. As shown in Table 3.1 there are five classes of strategy each seeking a different type of impression.

e 3.1

onomy of Self-Presentation Strategies Classified Primarily by Attribution Sought

	attribution sought	negative attributions risked	emotion to be aroused	prototypical actions
ingratiation	likeable	sycophant, conformist, obsequious	affection	self-characterisation opinion conformity, other enhancement, favours
intimidation	dangerous, (ruthless, volatile)	blusterer, wishy-washy, ineffectual	fear	threats, anger (incipient), breakdown (incipient)
self-promotion	competent, effective, ("a winner")	fraudulent, conceited, defensive	respect (awe, deference)	performance claims performance accounts, performances
self-implication	worthy (suffers, dedicated)	hypocrite, sanctimonious, exploitative	guilt (shame emulation)	self-denial, helping, militancy for a cause
self-implication	helpless (handicapped unfortunate)	stigmatised, lazy, demanding	nurturance (obligation)	self-depreciation, entreaties for help

2. From "Toward a General Theory of Strategic Self-Presentation" by E. E. Jones and T. S. Pittman (1982). In J. Suls (Ed.), Psychological Perspectives on the Self (Vol. 1). Hillsdale, New Jersey: Lawrence Erlbaum.

Ingratiation, as one class of self-presentation strategy, is aimed at intentionally deceiving the perceiver. The particular form of ingratiation, such as conformity, is determined by the ingratiator's resources and the situation (Jones & Baumeister, 1976). For example, it is more appropriate for a job applicant to engage in self-enhancing rather than conforming behaviour. Jones and Wortman (1973) feel that putting someone in a position of dependence prompts ingratiation behaviours. However, the importance of being liked by the particular target, the subjective probability of

success, the perceived legitimacy of the ingratiation behaviour, and individual difference factors all play a part in determining whether ingratiation will occur.

Intimidation is an attempt to extend interpersonal power by creating fears of negative consequences such as a child threatening to "make a scene", a manager who loses his temper, and a union official who readily resorts to industrial action. Jones and Pittman (1982) suggest intimidation is most likely in situations where the relationship between the target and the actor is nonvoluntary, the intimidator has ready access to negative consequences, the target person has poor retaliatory potential, and the intimidator is willing to forgo affection and other similar attributions.

Self-promotion as a self-presentation strategy seeks to create an impression of competence. This approach is illustrated by applicants being interviewed for jobs or training programmes, rugby players playing in a trial match for a representative team, or consultants discussing possible contracts with a client. In many cases the self-promoter has to cope with other sources of competency information, although often the response to such a possibility is to use indirect ways of creating the desired impression (Jones & Pittman, 1982).

Exemplification is aimed at creating the impression of integrity and moral worthiness by eliciting potential guilt in a target person who in turn is motivated to emulate the exemplifier. A "true" exemplifier is probably rare although such a strategy is frequently used by people such as parents, politicians, and certain political activists. Few such "worthy" persons are able to resist exploiting the power inherent in their own virtue (Jones & Pittman, 1982).

Supplication as the fifth self-presentation strategy stresses an inability for self-care and a dependence on others so making apparent an obligation or social responsibility (Berkowitz & Daniels, 1963). For example, a child may exaggerate his inability to complete various tasks around the home, a student may make obvious her impoverished state to an affluent friend, or a husband may avoid learning to sew. In each instance the supplicant is encouraging the assistance of others by acting as if helpless.

While it is conceivable that any of the above strategies could be used, self-promotion seems particularly relevant to applicants in personnel selection situations. To be seen as competent in respect of a particular position would seem to be central to the purpose of applying for a job. One problem for the job applicant engaging in self-promotion is that other measures of his or her competence are often available. It has been shown that people manage their self-presentations so as not to contradict negative information as well as being suitably modest in the presence of positive information (Baumeister & Jones, 1978). In addition and particularly in the presence of other diagnostic information, self-promoters encourage attributions of competence which exist for the most admired reasons (Quattrone & Jones, 1978). For instance, an applicant for a place in an orchestra may emphasise her natural ability (rather than many hours of practice), a reason that would seem to facilitate the likelihood of being a brilliant orchestral musician. "Most of us would rather be considered as relaxed but brilliant, for example, rather than plodding over-achievers" (Jones & Pittman, 1982, p. 259).

A further problem for self-promoters is what Jones and Pittman (1982) term the "self-promoter's paradox". We learn that many people exaggerate their abilities so their competency claims can be partially discounted. The paradox arises because claims are more likely when competence is doubtful than when it is certain (e.g., Wicklund &

Gollwitzer, 1983). When individuals make claims of superior ability, which are based on credible evidence, the attribution of competence may be obtained along with attributions of arrogance, insecurity, or even conceit.

Mechanism to Agency: Reciprocal Determinism and Self-Efficacy

The second major trend occurring in the area of self research is the movement from a mechanistic view of the self as a product of environmental influence, to a description of the self as an agent. Bandura's (1978) notion of reciprocal determinism is a approach which is explicit, both in its rejection of unidirectional or bi-directional determinants of behaviour, and in its treatment of the self as an agent.

Social learning theory (Bandura, 1977a, 1978, 1982a) places the self in a causal model of psychological functioning based on triadic reciprocity. Behavioural factors, environmental influences, and cognitive and other personal factors are seen to operate as interlocking determinants of each other (Bandura, 1977b). This view stands in contrast to a unidirectional view of interaction which treats persons and situations as independent entities but ignores reciprocal effects such as the influence individuals have on the environment through their behaviour. Similarly, a partially bi-directional view, while acknowledging the interaction of persons and the environment, treats behaviour as a byproduct. The effect of behaviour on the environment or a person's feelings through, say, playing the piano or switching on a television set is ignored (Bandura, 1978).

Bandura (1982b) describes a number of ways in which self processes enter interactively into the determination of behaviour. For example, personal factors affect how people select and transform environmental stimuli. Similarly, cognitive

capabilities guide behaviour by symbolically testing alternative courses of action and generating inducements to motivate goal directed behaviour.

Self-regulation involves three component processes; observation, judgment, and reaction (Bandura, 1982b). Firstly, to affect the direction of their own actions people have to observe their behaviour and its effects. From the large amount of information available, what is actually focused on depends on functional relevance and the individual's value orientation. Self-observation may be intermittent or regular, emphasise ongoing behaviour, or concentrate on the more distal results of behaviour. In addition, performance can vary on dimensions such as rate, quantity, originality, morality, etc. (Bandura, 1977b).

A second component of self-regulation is judgmental reaction. This involves contrasting what is observed with personal and referential information (e.g., reference group norms), assessing the value or relevance of the activity, as well as deciding whether or not performance can be attributed to one's own abilities and efforts.

The third component is self-response based on prior observation and judgment. Individuals create incentives for their own actions and apply self-consequences based on how their behaviour measures up to internal standards. We pursue courses of action that result in positive self-reactions and avoid behaving in ways that give rise to self-censure (Bandura, 1982b). Self-incentives provide guides and proximal motivators for courses of action often in combination with more distal incentives for the results of behaviour.

Self-efficacy judgments play a protective role when misjudgments of personal efficacy in either direction can have negative consequences. Gross overestimation of one's

capabilities may result in the undertaking of tasks that are beyond one's ability so resulting in unnecessary failure or even injury. Underestimation may mean individuals avoid beneficial environments and activities or focus on personal shortcomings generating stress that creates internal obstacles to effective performance (Bandura, 1982b). Some behaviours which are habitual do not require efficacy judgments as self-knowledge has been developed from repeated experience. Also, in many situations, reassessment is not necessary unless there is some significant change in a task or situation (Bandura 1982a).

Self-efficacy judgments also have motivational properties. When individuals commit themselves to specific goals, perceived differences between what they do and what they seek to achieve creates internal dissatisfaction which acts as an incentive for intensified effort (Bandura, 1978). However, in the absence of goals or information feedback regarding performance, cognitive comparison does not appear to enhance performance (Bandura & Cervone, 1983).

Bandura (1977b) indicates that self-efficacy judgments are principally based on four sources of information: enactive attainments, vicarious experiences, verbal persuasion, and physiological state. Enactive attainment is regarded as the most powerful source as it is based on instances of personal mastery. Successful performance raises expectations of mastery and repeated failures lower expectations, especially if such failures occur early in the learning process. The impact of enactive attainments depends on factors such as task difficulty, level of effort, situational aids, and the pattern and rate of success. Strong efficacy expectations developed through mastery performance generalise to other situations, even when performance is impaired by preoccupation with personal inadequacies (Bandura, Adams, & Beyer, 1977).

Performance desensitisation is one example where the power of enactive experience is

used as a treatment (e.g., Sherman, 1972). Work sample experience has also been suggested as having an effect on self-assessment of personal capabilities in a personnel selection situation (e.g., Downs, Farr, & Colbeck, 1978) and in choice of trade training (Downs, 1984).

Vicarious experience is simply experience gained through watching others perform. Vicarious information can influence efficacy appraisals both positively and negatively (e.g., Brown & Inouye, 1978). Such information can generate weak expectations, however, especially when observers have had little previous experience on which to base evaluations of their own competence (Bandura, 1982b). Generally, self-efficacy expectations induced by modelling are weaker than those derived from performance accomplishments as they are seen as less dependable (Brief & Aldag, 1981).

Verbal persuasion as a way of influencing human behaviour induces efficacy expectations that are likely to be weaker than those based on one's enactive experience (Bandura, 1977b). For example, a long history of a person being told they would never be able to drive a car may induce an efficacy-based expectation "unable to drive" that could easily be changed by one successful attempt at the controls of a vehicle. When verbal feedback is consistent with one's expectancies it is more accurately retained, given more credibility, and felt to result more from one's own abilities than feedback which is inconsistent with initial perceptions (e.g., Shrauger, 1975).

Emotional arousal can also provide information about self-competency. Individuals rely partly on their state of physiological arousal in judging their anxiety and susceptibility to stress. Bandura (1977b) suggests that high arousal impairs

performance so individuals are more likely to expect success when they are not tense and anxious.

The conception of agency as depicted by Bandura's (1982) reciprocal interactionism ascribes an important role to environmental factors. Self-regulatory mechanisms are dependent on learning processes which derive information from a variety of environmental sources. Gergen (1984) feels that agency in this context ultimately becomes a pawn to environmental contingency. This may be a rather harsh judgment on the self as an agent. Although environmental factors are undoubtedly important, especially in the initial stages of an individual's development (e.g., Bem, 1972), reciprocal interactionism states that cognitive, environmental, and behavioural factors can all take the role of initiator in self-regulation (Bandura, 1982a).

Self-Knowledge to Social Construction: Social Comparison

A third important trend in self research is the increasing emphasis on treating the self as a reflection of social context. The early focus in self research was centred on individual knowledge derived from internal states and observation of one's overt behaviour (Gergen, 1984). More recently the importance of the social context has been emphasised (Markus & Wurf, 1987). That knowledge of the self is a product of social interaction is illustrated in the continuing line of research on social comparison.

The theory of social comparison processes as discussed by Festinger (1954) is based on the assumption that within an individual there exists a drive to evaluate his or her own opinions and abilities. In most situations objective, nonsocial information which an individual can use to verify an opinion or assessment is not available. Information

is therefore sought by comparing one's own opinion or ability with that of others.

The theory also shows that we are more likely to choose someone close to our own ability or opinion for comparison. In the absence of a physical and social comparison, evaluations of abilities and opinions are likely to be unstable. According to Festinger, self-estimates of ability are subject to biases which do not affect opinions. For instance, higher performance is more valued than lower performance. In addition, ability estimates are likely to be more stable because opinions can be changed whereas abilities themselves cannot.

According to Sanders and Mullen (1984), the limited research on the impact of social comparison information on self-evaluation of specific abilities makes two important points. Firstly, the impact of performance comparisons on estimates of ability increases as the general similarity between comparison sources and evaluator increases. Secondly, estimates of ability are more influenced by modal scores than by extreme scores in a distribution of comparison information. In a study of the impact of comparison information Sanders and Mullen (1984) found that ability evaluations were largely free of distortions or bias. They concluded that people typically operate in a highly rational way only exhibiting distortion and bias when it is easier or less costly than logical information processing. What factors make rationality a costly option obviously requires further research (Sanders, 1985; Vries & van Kippenberg, 1985).

The use of social information does not necessarily increase objective accuracy as there is nothing in the scanning of social information that ensures an objective outcome. One's self-evaluation could be just as dependent on one's own ability as who is used for comparison and when and where the particular performance being compared took place (e.g., Hoffman, Festinger, & Lawrence, 1954; Morse & Gergen, 1970; Suls & Mullen, 1982).

The objective accuracy of judgments based on social comparison is also questioned on the grounds that much of the information used is based on inference and not direct observation (Gergen, 1984). Trying to assess another's ability involves an attributional problem because an ability is not necessarily directly observable. Whether an observable performance reflects ability, mood, motivation, physical stamina, or a number of other possibilities has to be inferred. However, it does seem that encouraging individuals to attend to the performance of others has a positive effect on the accuracy of self-evaluation. Mabe and West (1982) showed that self-evaluations based on comparison with other similar individuals tended to be more accurate than when such social information was not used.

Self-Assessment and the Self

While there is much debate over how individuals use and store information regarding the self, our ability to know and assess ourselves is generally accepted (e.g., Bandura, 1978; Bem, 1972; Mischel, 1972; Wicklund, 1975). As Gergen (1984) has remarked, one of the most important types of self-experience is evaluative in nature. As discussed above, application of self-presentation techniques require self-assessment of abilities and resources. In addition, self-efficacy judgments are a crucial part of self-regulation of behaviour and social comparison is based on a drive to evaluate one's own opinions and abilities.

Trope (1983) proposes that we use self-assessment to predict outcomes when we are uncertain about such factors as our ability, the effect of our ability on a task, and the importance of task outcomes. He suggests self-assessment becomes particularly important in those situations which depend heavily on our personal ability such as

prolonged, heterogeneous, and comparative tasks like careers and academic courses.

Other approaches also regard self-evaluation as a constant and necessary part of our everyday lives. An example is the self-evaluation maintenance model (SEM) which is based on the assumption that individuals are motivated to maintain a positive self-evaluation of themselves. As stated by Tesser and Campbell (1983):

"they [individuals] want to believe that they are competent; they want others to believe that they are competent; and they will behave so as to maintain a positive self-evaluation. If a situation promises an inevitable loss in self-evaluation, they will behave so as to minimise the loss. If a situation promises an opportunity to increase self-evaluation, they will take that opportunity" (p. 5).

The dynamics of the model are reasonably complex and well portrayed by Tesser & Campbell (1983). An important component of the model is that self-definition is assumed to be determined socially. Information is gained, not only from one's own level of performance, but also from the performance of others. In addition, other people can bring gains and losses in self-evaluation. For instance, an individual can gain through self-reflection when a close associate is highly competent in a area which is not central to his or her strivings. Conversely, a social comparison process occurs when a person's own skills are brought into doubt by the presence of someone who is an accomplished performer in the same domain.

Self-assessment is considered by some to be of superior validity to similar assessments provided by others. Mischel (1968, 1972, 1984) states that people tend to be the best predictors of their own attributes and future actions. The commitment to indirect personality testing (e.g., clinical assessment, personality inventories) demonstrated in the literature is considered inappropriate as the assumed relationship between overt behaviours and underlying dispositions is blurred by hypothesised defences that distort

and disguise the true meaning of behaviour (Mischel, 1972). Specific examples, such as that provided by Peterson (1965), show that self-assessments can be at least as accurate and more reliable than cumbersome and complex personality rating schedules. Scott and Johnson's (1972) review of personality measures found that direct (self-report) methods were generally favoured over indirect (disguised) approaches. Three studies conducted by the same authors found "no evidence for the superior validity of any of the indirect measures of personality characteristics over the direct methods against which they were compared" (p. 317). McClelland (1972) however disagreed saying that direct measures were only opinions predicting opinions.

The evidence supporting direct self-assessment measures is not confined to the field of personality. Shrauger and Osberg (1981) incorporated a wide range of criteria in a comparative review and concluded that self-assessment is favoured over other forms of prediction. Other reviews have also supported the validity of self-assessment (Mabe & West, 1982).

Such a positive approach to the value of self-reported evaluation is not a universal one. One criticism is based on symbolic self-completion theory. According to Wicklund and Gollwitzer (1983) this viewpoint assumes that important flaws in a person's training or performance are disguised by self-symbolising efforts such as the use of open self-aggrandisement and the display of status symbols. These efforts to influence others are considered signs of insecurity or incompleteness in the area expertise is sought. Indicators or symbols of completeness can be substituted for each other so that different sets of symbolic indicators (education, experience, etc.) can support the same self-definition aspired to.

Self-completion theory has direct implications for self-report validity. Incompleteness results in self-symbolising efforts for those committed to a certain self-definition. If, for instance, a job applicant feels incomplete in respect of an area of ability or expertise required in the position sought, self-reports are likely to be used to cover or compensate for the weakness. The tendency to use self-reports in a compensatory fashion results in so-called "antivalidity" (Wicklund & Gollwitzer, 1983). Experimental studies also show that when symbolic support (e.g., experience or qualifications) is lacking for a self-definition sought by an individual, self-report validity suffers (Wicklund & Gollwitzer, 1981). It must be noted that these effects are said to occur within the context of a single self-definition not a global or multi-dimensional self-view as suggested by Tesser and Campbell (1983).

Wicklund and Gollwitzer (1983) suggest that the relationship between what individuals say they will do and what they actual do should be a strong one but in many instances the association between prediction and behaviour is at best weak. Although judgments of one's own ability are functionally related to actions such as choice of activities, effort, and persistence, their accuracy may depend both on the self-motives being served (e.g., self-enhancement, consistency maintenance, or self-actualisation) and on the immediate social situation (Markus & Wurf, 1987). Making judgments about oneself is rarely a dispassionate act, particularly in regard to behaviour which has social importance (e.g., Anderson, Warner, & Spencer, 1984; Bauman & Dent, 1982).

Greenwald (1980) describes the self as a personal historian gathering and preserving an autobiographical picture which is prone to bias due to the self-referent nature of the information. These biases are termed egocentricity, benefectance, and cognitive conservatism. Individuals tend to best remember information that is highly relevant to the self as well as overestimate their own importance as an influence or target of

social interactions (egocentricity). People readily see themselves as responsible for positive outcomes and tend to deny responsibility for negative outcomes (beneffectance). Individuals tend to seek information that confirms their theories about themselves and to revise their autobiographical memory so it accords with their current self-concept (cognitive conservatism). As shown in the self-presentation literature, there are many subtle ways we project or convey a definition of a situation to suit ourselves (Goffman, 1969).

According to Bandura (1982b), the relationship between perceived efficacy and performance can also be affected by biases in self-monitoring and personal confidence in efficacy judgments. On many tasks, performance varies in relation to one's usual level of attainment. This variation allows some choice as to what performances are observed and remembered. Selective attention can therefore result in under or overestimates of efficacy even though the information has been accurately processed. Feedback which purposefully enhances performance information has been shown to result in enhanced self-efficacy judgment and performance. The strength of the relationship between self-efficacy judgments and subsequent performance can also be affected by a lack of personal knowledge regarding performance requirements, unanticipated situational constraints, and disincentives to act on self-appraisals of ability (Bandura, et al., 1977).

It also recognised that differences between individuals may have a profound effect on the validity of self-assessment. As an example, Mabe and West's (1982) review identified intelligence, achievement status, and locus of control as possible moderators of the validity of self-evaluation. In general terms, this review concluded that accurate self-evaluation was consistently in favour of more intelligent subjects, individuals high in achievement status, and those with an internal locus of control.

Another individual characteristic which would be expected to affect self-reports is social desirability. Crowne and Marlowe (1964) propose that the tendency to respond to self-reports in a socially desirable manner could conflict with the requirement for accuracy. The effects of social desirability and a number of other individual characteristics on self-evaluation are discussed in Chapter Four.

This chapter shows that the extensive psychological history of self research is moving from a preoccupation with static structure, unidirectional effects, and the belief that self-knowledge is derived from internal states, to considering the self as part of an ongoing process, as an agent reflecting social context. These trends serve to highlight the active role the self has in the regulation of behaviour and the central importance of self-evaluation in the self-control of behaviour. That we as human beings have demonstrated the ability to know and assess ourselves can be accepted. However, there is some dispute over whether such self-assessment information can be accurately reported and used in a predictive role. What has not yet been fully considered is the empirical research investigating self-assessment validity. Chapter Four describes and discusses the application of self-assessment with a particular focus on the world of work.

CHAPTER FOUR

THE APPLICATION OF SELF-ASSESSMENT

Empirical research using self-assessment as a source of information has been reported for a wide range of applications (e.g., Hunter & Hunter, 1984; Mabe & West, 1982; Reilly & Chao, 1982; Thornton, 1980). Examples include vocational guidance and performance appraisal (Levine, 1980), the prediction of therapy outcome, adjustment following hospitalisation, and intellectual achievement (Shrauger & Osberg, 1981).

Empirical Research Using Self-Assessment

Research has varied widely in the way self-assessments are requested and the type of criteria used. The form and content of self-assessment requests have ranged from direct self-evaluations of ability or performance (e.g., Shapiro & Dessler, 1985; Tharenou & Harker, 1984) to self-report measures which are only indirectly related to criteria used. Examples of such indirect measures include the Bills Index of Adjustment and Values (Bills, Vance, & McLean, 1951), a version of the California F Scale (Adorno, Frenkel-Brunswick, Levinson, & Sanford, 1950) and the Baron Ego Strength Scale (e.g., Dicken, 1969; Grande, 1966; Mischel, 1965). Other studies have used fictional autobiographies of the future (e.g., Brewster-Smith, 1966; Ezekiel, 1968; Tullar & Barrett, 1976) and "accomplishment inventories" as descriptions of past performance (Hough, 1984). Self-assessments have been related to training course performance (Harris, 1972), industrial ability tests (DeNisi & Shaw, 1977), typing tests (Anderson, Warner, & Spencer, 1984; Ash, 1980), and second language ability measures (Smith & Baldauf, 1982), as well as to various clerical tests (Levine, Flory, & Ash, [study 2], 1977), language, and arithmetical ability tests (Primoff, 1980).

In recent times self-assessment has been the focus of two extensive reviews, each approaching the topic in a slightly different way. Firstly, Shrauger and Osberg (1981) considered only those studies which compared self-assessment with other assessment procedures. Their findings showed that self-assessment was a better predictor of vocation and was more strongly related to peer assessment than were other methods such as personality tests. When compared to standardised tests and academic performance from previous settings, self-assessment was a better predictor of academic grades. However, when compared to previous performance in the same setting self-assessment was not superior. Reaction to therapeutic intervention was predicted marginally better by self, rather than other assessment methods, while the comparison of methods for assessing post-hospital adjustment was inconclusive. When all studies across all settings were combined, a significantly greater number of comparisons favoured self-assessment. Shrauger and Osberg (1981) concluded that other assessment methods did not show clearly superior validity when compared to self-assessment.

Mabe and West (1982) reported the second extensive review. They focused on the effects of individual characteristics and measurement conditions on the relationship between various ability and performance measures and self-assessment. Potential individual characteristics associated with accurate self-evaluation included relatively higher intelligence, high achievement status, and internal locus of control. The authors used meta-analytic techniques to show that when certain favourable conditions of measurement were present there was a substantial increase in the mean validity coefficient. The increase from .00 (no measurement conditions present), to .64 (any four measurement conditions) implied that many of the findings reported in the literature may be affected by the way in which self-assessment data were collected.

The four major contributors to explained variance in Mabe and West's (1982) analysis were relative self-evaluation, expectations of validation, anonymity instructions, and evaluation experience. Firstly, the relative self-evaluation measurement condition is based on Festinger's (1954) theory of social comparison. As discussed in Chapter Three, this theory proposes that if objective ability information is not available, individuals self-evaluate by comparing themselves with other people. As there are few physical, nonsocial sources of information, ability is better assessed by a comparison of one's ability or performance with that of others.

The second measurement condition, expectation of validation, is based on an individual's desire to obtain accurate self-evaluation information. According to social comparison theory, two motives operate in the self-evaluation of ability, the desire to obtain accurate information and the desire to enhance others perceptions of one's competence (Festinger, 1954). Mabe and West (1982) argue that enhancing the need for accurate information will positively influence the validity of the self-provided information. Such a condition can be implemented by telling raters that their self-assessment information is to be compared to criterion data.

The third measurement condition was a guarantee of anonymity to raters. This was expected to reduce the tendency to enhance self-assessments as such evaluations are not "publicised" or directly associated with the assessor. The risks connected with reporting are reduced and more valid reports are given. Anonymity and the application of self-assessment as a selection tool are not compatible. It would be ethically unacceptable to guarantee anonymity and then use the self-assessment information to assist in a selection decision.

The fourth condition of measurement, self-evaluation experience, is based on the premise that self-evaluation is a skill, one that could be improved by practice. Self-raters become more accurate if they are given an opportunity to practice, although it seems that such improvement would depend on feedback. Mabe and West (1982) concluded that as 76% of the correlations used in their analysis contained three or fewer measurement conditions it was not surprising that many of the validity correlations were low.

One other measurement condition that did not demonstrate a significant relationship with the strength of validity coefficients in Mabe and West's (1982) study was match, or similarity between the self-evaluation and the criterion measure. These authors, however, felt that the small variances on the match condition in the sample of studies used may have resulted in restriction of range so reducing the coefficient. The principle of a point to point relationship between predictor and criterion has been strongly supported by advocates of the work sample test (e.g., Asher & Sciarrino, 1974) and the assessment centre (Thornton & Byham, 1982).

Self-Assessment in Educational Research

The major area outside of work that has made use of self-assessment is the prediction of academic or intellectual achievement. An influential model in educational research dealing with the self-concept is that offered by Shavelson and colleagues (Shavelson & Bolus, 1982; Shavelson, Hubner, & Stanton, 1976) which has formed the basis of the Self Description Questionnaire (SDQ). Shavelson sees the self-concept as an individual's perception of self formed through environmental experience, interactions with significant others, and attributions about one's own behaviour. One's self-concept is both descriptive and self-evaluative. Individuals have a number of self-concepts that involve inference in academic and nonacademic areas.

A recent reformulation of Shavelson's model describes the multifaceted and hierarchical nature of the self-concept (Marsh & Shavelson, 1985). The general self-concept is split into subordinate self-concepts titled "Nonacademic", "Academic English", and "Academic Mathematic" which in turn rest on self-concepts such as "physical ability", "physical appearance", "peer relationships", "parent relationships", "reading", "general school", and "mathematics". Research using the SDQ has generally confirmed this structure (Byrne & Shavelson, 1986). Much of the work using Shavelson's model focuses on the structure of the self-concepts as well as bias involved in academic attributions and is based on research using preadolescent and adolescent children (e.g., Marsh, 1986; Marsh, Barnes, Cairns, & Tidman, 1984; Marsh, Smith, & Barnes, 1985).

In general terms, self-assessments of academic ability are positively and significantly correlated with measures of academic ability such as teacher ratings and test scores of ability. From a representative sample of correlational and experimental research, Byrne (1984) found the relationship between self-concept and academic ability was almost always positive and was strongest in matching content areas. The correlations between general academic self-concept and academic achievement measures were described as moderate. Nonacademic self-concept and academic achievement were generally uncorrelated. Byrne (1984) was unable to make any conclusions as far as a causal relationship between self-concept and academic ability was concerned.

Hansford and Hattie (1982) found that measures of ability/performance correlated around .2 with measures of general self-concept and .4 with measures of academic self-concept. These authors suggested that grade level, socioeconomic status, ethnicity, ability level, specificity of measures, and type of criteria may affect the

relationship between self-concept and academic achievement. Shavelson and Bolus (1982) have suggested other possible moderators including peer and parental influence. Research involving subject samples in tertiary educational settings generally report similar significant positive relationships between self-assessed competence and academic performance (e.g., Foot & Smith, 1986; Gadzella, Cochran, Parham, & Fournet, 1976; Moreland, Miller, & Laucka, 1981; Wheeler & Knoop, 1982).

The provision of self-assessment information in educational settings appears to have few implications for the assessor. Personnel selection and performance appraisal decisions which use self-assessments may have major implications for the individual over which the assessor has little control. The information provided by such a judgment may affect a promotion or a job offer. Self-assessments in educational settings, however, are collected generally for research purposes and have little or no obvious effect on the assessor's academic progress, although there are exceptions (e.g., Henbest & Fehrsen, 1985).

Self-Assessment in Vocational Guidance

Self-assessment information is used extensively in vocational guidance as part of a collaborative process to assist in career decision making. Overall, expressed interests tend to be better predictors of vocational choice than interest inventories (Shrauger & Osberg, 1981), a conclusion compatible with earlier reviews of this field (e.g., Whitney, 1969). Self-evaluation of skill strengths and weaknesses, preferences, values, and strategies are, along with searches for career information, major components of the career decision making process (e.g., Krumboltz, Mitchell, & Jones, 1976; Noe & Steefy, 1987). The extensive use of self-assessment information in this area not only reflects the importance of such information but also assumes individuals are accurate assessors of themselves. This assumption is not always supported, however, as self-

assessments of abilities, skills, and aptitudes are often strongly influenced by interests and biases due to a lack of realism, trait confusion, and heightened stereotypy (Lunneborg, 1982). Whether vocational guidance self-ratings suffer from positive leniency effects similar to those shown in personnel selection situations has not been established. One study focusing on the accuracy of self-ratings for career decision making did report positive leniency effects but used an assessment centre to provide assessor ratings in a situation that seemed little different from a laboratory study of personnel selection (Mihal & Graumenz, 1984).

An interesting line of research in vocational guidance is the investigation of the effect of self-efficacy expectations on the career choice of men and women. Hackett and Betz (1981) describe the failure of women to utilise their individual capacities, talents, and interests in career pursuits as largely due to socialisation experiences. Research does show that although female job preferences and aspirations are similar to males, their range of occupation choice tends to be much lower (e.g., Hoult & Smith, 1978). Individuals will choose and find most satisfying those roles which are consistent with how they see themselves (Korman, 1970).

Bandura (1977b) suggests behaviour and behaviour change are mediated primarily by expectations of personal efficacy. When the sources of self-efficacy information are considered, the differential effect of enactive, vicarious, verbal persuasion, and emotional arousal information on males and females is rather obvious. For example, role models in a range of career options are largely absent for females resulting in a loss of vicarious information. Similarly, females have a greater enactive experience in domestic and nurturance activities and less in mechanical, sporting, and other traditional "masculine" domains (Deaux, 1979; Hackett & Betz, 1981). Compared to males, females report significantly higher levels of self-efficacy in respect of

traditionally "feminine" occupations and significantly lower levels of self-efficacy for nontraditional or masculine occupations (Betz & Hackett, 1981). It seems that females may actually avoid pursuing occupations of which they are capable due to beliefs in personal ability.

The Validity of Self-Assessment in Work and Work-Related Situations

A review of the literature dealing with self-assessment in work and work-related applications was conducted to summarise validity-related information and to illustrate the variety and extent of research in this area. Studies were included if self-assessment of one's own ability or skill on some dimension related to work formed part of the research. Investigations with no direct or specified link with work were excluded (e.g., Shrauger & Patterson, 1974; Smith & Baldauf, 1982; Toughey, 1972). In addition, some studies which used self-assessment measures could not be included as they did not report self-assessment/criterion relationships in any form (e.g., Bassett & Meyer, 1968; Chiu, 1975).

Although self-assessment has been used as a source of predictor information in work-related areas for over 30 years, a large portion of the research often is not readily available appearing as unpublished dissertations, reports with very limited circulation, personal communications, or as incidental elements of other research reports.

To summarise the predictive validity demonstrated by this review a weighted average of reported correlations was used as a best estimate of the population correlation with the variance being described using a frequency weighted average squared error, s^2 (Hunter, Schmidt & Jackson, 1982). In all, 46 reports made up the review, four of which were not included in summary statistics as data equivalent to correlations were

not reported (Farley & Mayfield, 1976; Schendel & Hagman, 1982; Shapiro & Dessler, 1985; Shore & Thornton, 1986). A number of papers included more than one sample. Each independent sample and its associated correlation was treated as a separate study as recommended by Hunter et al., (1982). A brief description of each report, along with corresponding validity data, is shown in Table 4.1.

From the 46 studies, a total of 53 correlation coefficients were examined.

Correlations ranged from $-.07$ to $.81$ with a population correlation of $.18$ ($s^2 .0329$).

This is slightly higher than the average weighted correlation of $.15$ based on three coefficients reported by Reilly and Chao (1982) and substantially higher than $.00$ based on four coefficients given by Hunter and Hunter (1984). Mabe and West's (1982) review reported a weighted population correlation of $.19$ based on 267 separate samples involving self-assessment of work performance, scholastic achievement, cognitive abilities, and various other skills.

The set of studies used in the present review includes one particularly large sample of 8493 (Bass & Burger, 1979). Bass and Burger's (1979) research was a cross-cultural investigation using self-ratings of risk tolerance, interpersonal competencies, and managerial style made by male managers during management training workshops. The average correlation of self-assessments with a rate of advancement index, as computed by Hunter and Hunter (1984), was a low $.05$ with no variation after sampling error was eliminated.

Table 4.1: Application of Self-Assessment as a Predictor in Work and Work-Related Situations

STUDY	PREDICTOR	SAMPLE	METHOD	CRITERIA	RESULTS & COMMENTS
Parker, Taylor, Barrett & Martens, 1959	Direct measure; job dimensions	Clerical workers n = 117	Concurrent	Supervisor performance ratings	Average correlation of 7 job dimensions, $r = .32$ ($p < .01$) SE .061. Overall job performance, $r = .35$ ($p < .01$).
Kriedt & Dawson, 1961	Indirect measure; Personal Inventory	Clerical workers n = 41	Concurrent	Supervisor performance ratings	Total Personal Inventory score correlated .47 ($p < .01$) with supervisor ratings.
Mayo & Manning, 1961	Indirect measures; self-rating of training course effort	Aviation structural mechanic trainees (navy) n = 196	Concurrent & Predictive	<u>Concurrent</u> aptitude tests of cognitive ability (GCT), math (ARI), mechanical (MECH), peer rating, behavioural checklist (motivation), school grades; <u>Predictive</u> final course grades in sheet metal & welding	<u>Concurrent</u> self-ratings with GCT -.09 (NS), with ARI = -.10 (NS), with MECH .02 (NS), with peer rating .28 ($p < .01$), with motivation checklist -.03 (NS), with school grades .09 (NS). Average of concurrent measures .03 (NS) SE .024. <u>Predictive</u> self-ratings of effort with final course grades in sheet metal $r = .19$ ($p < .01$), with welding $r = .04$ (NS).
Prien & Liske, 1962	Direct measures; self-ratings of 8 dimensions of work performance	Variety of operators n = 96	Concurrent	Supervisor ratings of performance on 8 dimensions of work	The average reported correlation between 1st level supervisor and self-ratings was .25 ($p < .05$). Self/2nd level supervisor also reported ($r = .13$, NS)
Kirchner, 1965	Direct measures; job dimensions & overall job performance	Technical personnel n = 92	Concurrent	Supervisor performance ratings	Average of 4 dimensions of work, $r = .23$ ($p < .05$) SE .133. Overall job performance, $r = .22$ ($p < .05$).
Mischel, 1965	Indirect measures; Ego Strength Scale (Es), California F Scale (F), Manifest Anxiety Scale (MA)	Peace Corp volunteers n = 41	Concurrent & Predictive	<u>Concurrent</u> academic performance, peer ratings, training staff, interview & board ratings; <u>Predictive</u> field staff ratings	<u>Concurrent</u> no self-report measures were significantly related. Average correlation .02 (NS) SE .036. <u>Predictive</u> Es, .34 ($p < .05$) F -.45 ($p < .01$), MA, -.34 ($p < .05$). Self-report measures were better predictive measures in all cases.
Pym & Auld, 1965 (3 samples)	Direct measures; <u>All samples</u> , Self ratings of competence	<u>Sample 1</u> machinists n = 85; <u>Sample 2</u> commerce apprentices n = 35; <u>Sample 3</u> Engineers n = 89	Concurrent	<u>Sample 1</u> ratings of competence and overall performance; <u>Sample 2</u> examinations, work reports, & supervisor ratings; <u>Sample 3</u> 3 ratings of competence (technical, customer relations, and cooperation)	<u>Sample 1</u> correlations for foreman's competence rating, .52 ($p < .001$), foreman's overall assessment, .42 ($p < .001$), averaged over 5 output periods $r = .50$ ($p < .001$) SE .005. Over three criteria average $r = .48$ ($p < .001$) SE .030. <u>Sample 2</u> correlations for exams .25 (NS), work reports .46 ($p < .05$), supervisor ratings average $r = .63$ ($p < .05$). Average r over all criteria .46 ($p < .05$) SE .118. <u>Sample 3</u> average self/supervisor rating $r = .59$ ($p < .001$) SE .055.
Grande, 1966	Indirect measure; Bills Index Adjustment and Values (B)	Peace Corp volunteers n = 62	Predictive	Field staff rating	Field ratings were used to group volunteers into superior (17), average (26) and ineffective (19). No significant differences found between extreme groups.

STUDY	PREDICTOR	SAMPLE	METHOD	CRITERIA	RESULTS & COMMENTS
Lawler, 1967	Direct measures; Ratings of quality of performance, ability & effort on the job	Middle and top level managers n = 113	Concurrent	Superior and peer ratings of job performance	Self/superior ratings for quality (.01), ability (.13) and effort (.30) gave an average r of .15 (NS) SE .084. Self/peer ratings for quality (.01), ability (.09) and effort (.30) gave an average of .13 (NS) SE .086.
Thornton, 1968	Direct measures; self-appraisal of job performance (27 items)	Upper level managers n = 64	Concurrent	Superior performance appraisals and 'promotability index' (PI)	The average correlation between the 27 items used in the self-/superior ratings of job performance was .22 (NS) not .23 as reported SE .030. Self-ratings/PI average r = -.04 (NS) SE .020. Supervisor/PI was .15 (NS) SE .026. Strong leniency effects.
Dicken, 1969	Indirect measure; Baron Ego Strength Scale (Es)	Peace Corp volunteers n = 53	Concurrent & Predictive	<u>Concurrent</u> final selection rating; <u>Predictive</u> composite field rating	<u>Concurrent</u> Es correlated .22 (NS) with board rating. <u>Predictive</u> Es correlated .32 (p<.05) with field rating.
Touhey, 1971	Direct measures; self-estimate of clerical ordering task	Undergraduate students n = 46	Concurrent	Clerical, verbal and arithmetic tasks	Self-estimates were not significantly related to clerical task scores. Verbal and arithmetic task correlations were .29 (p<.05) & .55 (p<.01) respectively. Average r = .29 (p<.05) SE .147.
Harris, 1972	Direct measures; performance dimensions (technical skills, motivation, interpersonal skills, language & adaption)	Peace Corp volunteers 2 studies n = 29 & n = 18	Concurrent	Mid-training and final training course rating	1st study (n = 29) only motivation (.48, p<.01) significant. Average r = .19 (NS) SE .055. 2nd study (n = 18), significant correlations were motivation (.61, p<.01), interpersonal skills (.45, p<.05) and adaption (.59, p<.01). Average r = .43 (p<.05) SE .054.
D'Augelli, 1973	Direct measures; group Assessment of Interpersonal Traits (GAIT), professional psychotherapy skills	Undergraduate students n = 168	Concurrent	Peer and observer rating	5 of the 8 GAIT dimensions were significantly related to equivalent peer ratings (average r = .22, p<.05 SE .041. No self/observer correlations were significant r = .08 (NS) SE .043.
Williams & Seiler, 1973	Direct measures; dimensional and global measures of work performance and effort	Engineers n = 202	Concurrent	Superior ratings of job performance and effort	Self/superior correlations for global ratings were .24 (p<.01) for effort and .48 (p<.01) for performance. The same ratings derived from a combination of 5 dimensions were .33 (p<.01) for effort and .60 (p<.01) for performance. The average of the two methods for performance was .54 (p<.01) SE .060.
Heneman, 1974	Direct measures; self-ratings on dimensions and overall job performance	Managers n = 102	Concurrent	Superior ratings on dimensions and overall job performance	Self/superior correlations for dimensions were planning .31 (p<.01), investigating .18 (NS), coordinating .17 (NS), evaluating .02 (NS), supervising .39 (p<.01), staffing .07 (NS), negotiating .32 (p<.01) and representing .27 (p<.05). Overall effectiveness r = .26 (p<.01). Self-ratings were less lenient, showed less halo and were more variable than superior ratings.
Klimoski & London, 1974	Direct Measures; self-ratings of job performance	Registered nurses n = 153	Concurrent	Supervisor and peer ratings of job performance	Self/supervisor rating .05 (NS). Self/peer rating .08 (NS).

STUDY	PREDICTOR	SAMPLE	METHOD	CRITERIA	RESULTS & COMMENTS
Ekpo-Ufot, 1976	Indirect measure; Self Perceived Abilities Related to the Task (SPART)	Automobile assemblers n = 123	Predictive	Turnover ('stays' & 'quits' after 6 months on the job)	A total SPART score was correlated .34 (p<.01) with turnover. The higher the SPART score, the more likely a worker was to stay on the job.
Farley & Mayfield, 1976	A blend of Direct and Indirect measures; Friends Test measuring a need to achieve and working hard plus a total score	Life insurance sales applicants n = 1119	Predictive	Turnover (1 year) and earned commission	The 'Friends' test was part of regular selection procedures. No significant relationship was found between either Friends test rating and either criterion score but a combined Friends test score was related to turnover (Chi square [4] = 11.28 p<.05).
Baird, 1977	Direct measures; self-assessment of job performance on five dimensions	A range of employees in a state agency (inc managers, technicians & clerical workers n = 165	Concurrent	Superior performance appraisal on ability, effort, quality, and overall performance	Self/superior ratings correlations for appraisal items, quality .12 (NS), quantity .21 (p<.01), effort .13 (p<.05), ability .16 (p<.05), overall performance .20 (p<.01). Leniency effect in self-assessments as well as less halo when compared to superior ratings. Relationships between ratings related to self-esteem and job satisfaction.
Brief, Aldag, & van Sell, 1977	Direct measures; self-appraisal of quality and quantity of work	Housekeeping and food service personnel in a hospital setting n = 129	Concurrent	Supervisor performance ratings	Self/supervisor correlations for quality and quantity of work not significant. Average r = .1 (NS) SE .020. No moderator effects for race, gender, age, satisfaction levels or education. Short tenure employees (2.2 yrs or less) were more accurate than long tenure employees.
DeNisi & Shaw, 1977	Direct measures; work related abilities (e.g. visual pursuit, manual speed and accuracy, general intelligence, etc.)	Undergraduate students n = 114	Concurrent	Ability test scores	Self/ability test correlations were visual pursuit .05 (NS), manual speed and accuracy .19 (p<.05), verbal ability with 2 tests .29 (p<.001) & .36 (p<.001), numerical ability with 2 tests .41 (p<.001) & .37 (p<.001), general intelligence with Otis .26 (p<.001) & SAT combined .35 (p<.001), spatial orientation .21 (p<.05) and mechanical comprehension .36 (p<.001). Average r = .29 (p<.001) SE .034.
King & Manaster, 1977	Direct measure; self-rating scale of job interview performance	Undergraduate students, female n = 98	Concurrent	Trained judge's ratings	Job interview expectations and self-ratings of interview performance were not significantly related to judges' ratings of interview performance (r = .22, NS).
Levine, Flory, & Ash, 1977	Direct measures; <u>Study 1</u> used self-ratings of clerical abilities. <u>Study 2</u> , self-rating of typing speed	<u>Study 1</u> clerical workers n = 73; <u>Study 2</u> clerical applicants n = 569	Concurrent (both 1 & 2)	<u>Study 1</u> clerical test scores, supervisory performance ratings; <u>Study 2</u> typing test speed score	<u>Study 1</u> self-assessment/sub-scale correlations; spelling .58 (p<.01), reading/comprehension .32 (p<.01), grammar .44 (p<.01), word meaning .47 (p<.01), proof reading .07 (NS), arithmetic .43 (p<.01), alphabetizing (NS), comparison (NS). Average r = .32 (p<.01) SE .0676. Seven correlations were significant. Average r = .22 (NS) SE .0524. <u>Study 2</u> sample divided into 3 cohorts. Average correlation was .64 (p<.01) SE .008.

STUDY	PREDICTOR	SAMPLE	METHOD	CRITERIA	RESULTS & COMMENTS
Bass & Burger, 1979	Indirect measures; risk tolerances, interpersonal competencies & managerial style	Middle managers, tertiary educated, male n = 8493	Concurrent	Rate of advancement score	No correlations of self-assessments with rate of advancement reported, although Hunter & Hunter, 1984 computed an average correlation of .05 (NS).
Ekpo-Ufot, 1979	Direct measure; self-perceived abilities relevant to the task (SPART)	Clerical workers n = 88	Concurrent	Supervisor performance ratings & complaining behaviour	SPART score was related to supervisor ratings $r = .23$ ($p < .05$) an number of complaining behaviours $r = -.30$ ($p < .05$). Gender as a moderator was not significant.
Ash, 1980	Direct measures; self-ratings of typing skills	'Job ready' high school students n = 156	Concurrent	Typing test scores (straight copy letters, revised manuscript, tables and numbers)	Self-assessment/test score correlations were straight copy, .59 ($p < .01$), letters .27 ($p < .01$), revised manuscript .27 ($p < .01$), tables .07 (NS), numbers .08 (NS). Self-assessment of straight copy a 'best predictor' of other skills (letters $r = .48$, revised manuscript .44, tables .26, numbers .30, all $p < .01$). Majority/minority moderator effect was not significant.
Primoff, 1980 (derived from 4 reports of research from previously unpublished or limited circulation reports with sufficient detail for this table)	Direct measures; <u>Report A</u> : clerical skills, <u>Report B</u> : equipment operation estimate, <u>Report C</u> : filing ability, <u>Report D</u> : facilities repair self-assessment.	<u>Report A</u> clerical applicants n = 1000 +; <u>Report B</u> heavy equipment operators n = 93; <u>Report C</u> 2 samples of clerks n = 558 & n = 1040; <u>Report D</u> facilities repairers n = 326	Concurrent	<u>Report A</u> test (spelling, word meaning, multiplication & comparison; <u>Report B</u> operational test; <u>Report C</u> tests; <u>Report D</u> promotion exam	<u>Report A</u> self-assessment/test score correlations, spelling .44 ($p < .01$), word meaning .50 ($p < .01$), multiplication .40 ($p < .01$). <u>Report B</u> : $r = .61$ ($p < .01$). <u>Report C</u> sample 1, $r = .33$ ($p < .01$), sample 2, $r = .34$ ($p < .01$). <u>Report D</u> : $r = .81$ ($p < .01$).
Kesselman, Lopez, & Lopez, 1982	Direct measures; self-scored in-basket on 33 dimensions	1st line supervisors n = 80	Concurrent	Assessor ratings of overall performance trait ratings (T) and behavioural observation report (BOR)	Correlations between self and T based ratings .33 ($p < .01$). Self and BOR based ratings .27 ($p < .05$).
Schendel & Hagman, 1982	Direct measure; estimate of number of retraining trials	Soldiers n = 42	Predictive	Disassembly/assembly of M60 machine gun	An average of 1.97 trials were needed to restore criterion performance. 73% of soldiers' estimates were within 1 trial of being correct. Error means reported suggest accurate assessments
Wheeler & Knoop, 1982	Direct measures; self-ratings of teaching performance using the 'Performance Improvement Guide' giving an overall score	Student teachers n = 47	Concurrent	Performance Improvement Guide (PIG), 7 categories completed field & academic supervisors	One significant PIG self-assessment field supervisor rating, learning climate .34 ($p < .05$), average $r = .14$ (NS) SE .040, overall performance rating .11 (NS). Three significant PIG self-assessment/academic supervisor ratings, student relation .26 ($p < .05$), relations with peers & principals .37 ($p < .01$) and professional qualities .29 ($p < .05$), average $r = .24$ (NS) SE .028 overall performance rating .32 ($p < .05$).

STUDY	PREDICTOR	SAMPLE	METHOD	CRITERIA	RESULTS & COMMENTS
Anderson, 1983	Direct measures; self-ratings of overall job performance	<u>Sample 1</u> middle managers male, white n = 44; <u>Sample 2</u> 1st level white, male n = 40; <u>Sample 3</u> 1st level, male Polynesian n = 19; <u>Sample 4</u> middle level, female, white n = 31	Concurrent (all samples)	Leadership effectiveness rating (all samples)	<u>Sample 1</u> r = .51 (p<.001). <u>Sample 2</u> r = -.03 (NS). <u>Sample 3</u> r = .08 (NS). <u>Sample 4</u> , r = -.07 (NS). Note in this study Polynesian refers to New Zealand Maori as well as Pacific Island immigrants.
Anderson, Warner, & Spencer, 1984 (Study 2)	Direct measures; self-ratings of typing speed	Applicants for clerical positions n = 66	Concurrent	Speeded typing performance test	Typing speed self-assessment and test correlated .27 (p<.05), corrected for inflation r = .48 (p<.001).
Mount, 1984	Direct measures; self-ratings of managerial performance	Middle level managers n = 80	Concurrent	Superior performance rating	3 of 8 managerial performance reported significant correlations (administration .56, p<.05, feedback .26, p<.05, consideration .35, p<.05). Average r = .18 (NS) SE .073.
Pannone, 1984	Direct measures; ratings of work experience on same dimensions as criteria	Applicants for electrician position n = 221	Concurrent	Electrical test performance	Self-rated experience significantly related to criteria (r = .42, p<.01). Applicants classified as fakers (n = 144) r = .26 (p<.05) and nonfakers (n = 75) r = .55 (p<.01).
Steel & Ovalle, 1984 <u>Study 1</u>	Direct measures; self-ratings of job dimensions	Branch managers of lending institutions n = 401	Concurrent	Superior performance ratings and branch performance using nett loan and account growth, financial control & profitability to form a total bonus score.	Compared conventional self-appraisal (SA) and self-appraisal based on how the rater's perception of supervisor feedback on performance (FBSA). Average SA ratings of quality, quantity, efficiency, problem solving, and adaptiveness correlated .19 (p<.001) SE .017 with overall supervisor rating and .08 (NS) SE .020 with total bonus score. FBSA ratings on the same dimensions, average of .35 (p<.001) SE .014 with supervisor rating and .10 (p<.05) SE .022 with total bonus score.
Steel & Ovalle, 1984 Study 2	Direct measures; self-ratings of job dimensions	Military staff; <u>Sample A</u> administration n = 108; <u>Sample B</u> maintenance personnel n = 142	Concurrent	Supervisor performance ratings of quality, quantity, efficiency, problem solving and adaptability	Both samples examined FBSA (see study 1) in conditions of high and low (LF) feedback on job performance. <u>Sample A</u> LF (n = 54) average of 5 dimensions r = .23 (NS) SE .029. HF (n = 51) average = .26 SE .081. <u>Sample B</u> LF (n = 74) average r = .18 (NS) SE .01 (n = 68), average r = .36 (p<.05) SE .049.
Tharenou & Harker, 1984	Direct measures; self-ratings of job performance	Electrical apprentices n = 166	Concurrent & Predictive	Supervisor performance rating and self-appraisal of job performance	Self and supervisory ratings were made at Time 1 and Time 2 (21 months later). Concurrent T1 (n = 166) r = .13 (NS), T2 (n = 92) r = .11 (NS). Predictive self T1 with supervisory T2 (n = 92) r = .20 (NS). NB Supervisory T1/T2, r = .30, p<.01).

STUDY	PREDICTOR	SAMPLE	METHOD	CRITERIA	RESULTS & COMMENTS
Arnold, Willoughby, & Calkins, 1985	Direct measures; self-ratings on average of 11 items of medical performance during training	Medical students n = 211	Concurrent & Predictive	Concurrent docent (tutor) ratings of medical performance; Predictive self-rating of medical performance during training	Concurrent self/docent ratings at 3rd, 4th, and 5th year of training $r^{(S)}$ = .25, .29, and .22 respectively (all $p < .01$). Self/Yr6 r = .10 (NS). Average concurrent r = .22 ($p < .01$). Predictive self/self ratings during training, Yr3/Yr4, r = .63, Yr3/Yr5, r = .60, Yr3/Yr6 r = .54, Yr4/Yr5, r = .75, Yr4/Yr6, r = .70. Yr5/Yr6, r = .67. All sig at $p < .01$. Average r = .65 ($p < .01$) SE .030. Self-ratings at admission (Ad) with self-ratings after training Ad/Yr5 r = .17 ($p < .05$), Ad/Yr4 r = .29 ($p < .01$), Ad/Yr5 r = .24 ($p < .01$), Ad/Yr6 r = .15 ($p < .05$). Average r = .21 ($p < .01$) SE .032. No moderator effect due to gender.
Henbest & Fehrsen, 1985	Direct measure; self-assessment of medical skills (training)	Final year medical students n = 19	Concurrent	Faculty ratings of medical skills	Self/faculty ratings correlated .74 ($p < .01$). Rated as a very positive experience by students. Almost no leniency effects.
Florde, Wollitzer, & Blossom, 1985	Direct measure; self-assessment questionnaire for patient educators (doctors)	Family practice residents n = 24	Concurrent	Instructor ratings of residents' proficiency	Self-assessment correlated .60 ($p < .01$) with instructor rating.
Shapiro & Dessler, 1985	Direct measure; Job performance	Hospital supervisors n = 146	Concurrent	Superior performance ratings	Matched pairs t-test revealed significant differences between self- and superior ratings ($p < .001$).
Schmitt, Ford, & Stults, 1986	Direct measures; self-ratings of assessment centre performance, 8 dimensions (post centre)	Applicants for 1st line supervisory positions n = 1693	Concurrent (pre & post measure)	Assessor ratings, post-assessment centre ratings	Post-centre/assessor dimension ratings average r = .11 (SD .054). Pre-/post assessment centre self-ratings; organizing & planning .47, analyzing .47, decision making .43, controlling .42, communication .44, interpersonal relations .44, influencing others .40, flexibility .50. Average r = .45 ($p < .01$) SE .011.
Shore & Thornton, 1986	Direct measures; self-appraisal of work performance (6 aspects)	Electronic manufacturing assemblers n = 70	Concurrent	Supervisor performance ratings on 8 dimensions	Self-/supervisor ratings, work quality $\text{Eta}^{\#}$ = .31, work quantity Eta .32, judgment Eta .47, initiative Eta .42, teamwork Eta .33, dependability Eta .42, average Eta = .38 SE .027.
Staley & Shockley-Zalabak, 1986	Direct measures; self-rating of management and communication skills	Female professional managers n = 122	Concurrent	Supervisor performance ratings on 15 skills	Only 3 (business writing r = .41, $p < .001$, oral presentations r = .45 $p < .001$, communications technology [computers] r = .52 $p < .001$) correlations were significant. Other less 'concrete' skills (e.g. interviewing, negotiating & listening) were not significantly related. Average correlation .18 (NS) SE .040. Over a number of different organisations (not reported).
McEnery & McEnery, 1987	Direct measures; self-appraisals of job performance	Management & professional hospital staff n = 200	Concurrent	Superior ratings of job performance on 29 dimensions	Data supplied by authors on 15 dimensions only. Average correlation .12 (NS) SE .078

SE Standard Error

To reflect the magnitude of effect of rating source in cases where F tests only were reported, F tests were converted to Eta. Eta reflects the difference in group mean scores between the different sources of rating. Correlations on the other hand show how closely a group of subjects' '1st' rating matches their '2nd' rating on average.

The effect of this outlier study substantially reduced the average correlation of indirect and concurrent studies. When this study is removed, the population correlation rose to .30 ($S^2 .0329$). Although there are no clear guidelines of how to deal with the effect of such an outlier (Hunter et al., 1982), the variety of organisations and cultures, as well as the simulated nature of the setting, makes this investigation unique in this review and consequently less comparable with other work related investigations.

The studies reviewed showed that self-assessment was used in a number of different ways. As these various approaches possibly could have some effect on reported validity coefficients, the studies were grouped according to the approaches taken. Population correlations were calculated for each group and are shown in Table 4.2 .

Firstly, as mentioned above, some investigations used self-assessment measures which were not directly related to the criterion with which they were compared. For instance, self-assessment information was requested on concepts such as a need to achieve (Farley & Mayfield, 1976) and ego strength (Dicken, 1969). These measures were classified as indirect as they used self-assessments of abilities or traits which were not directly measured by criteria (e.g., Bass & Burger, 1979; Mayo & Manning, 1961; Mischel, 1965). In comparison, other investigations used self-assessment as a direct assessment of a specific ability or skill, such as self-ratings of overall job performance (e.g., Heneman, 1974; Shapiro & Dessler, 1985) or dimensions of work (e.g., Wheeler & Knoop, 1982).

Table 4.2

Sub-Groupings of Self-Assessment Correlations

Sub-Group	n of Samples	Total N	Average n	Average Correlation	Weighted SD (s ²)
All studies	53	18379	346.77	.18	.0329
Less *	52	9886	190.12	.30	.0340
Direct only	47	9432	200.68	.30	.0344
Indirect only	6	8947	1491.17	.06	.0182
Less *	5	454	90.8	.21	.0204
Concurrent	52	18256	351.08	.18	.0333
Less *	51	9763	191.43	.29	.0353
Predictive	6	790	131.67	.28	.0149
Selection Predictive	3	290	96.67	.19	.0191
Selection situation Concurrent	11	5613	510.27	.34	.0397

* An identical sample of studies as above but with Bass and Burger (1979) omitted from the analysis.

According to the notion that matched predictor and criterion measures are more likely to be strongly related than measures that are not matched (Asher & Sciarrino, 1974), it could be expected that predictors directly measuring the criterion would demonstrate higher validity coefficients than predictors which were indirect measures. When population correlations were grouped separately for direct and indirect measures of self-assessment, the direct measures demonstrated stronger relationships on average even when the large outlier study was removed from the indirect group (Table 4.2).

The population correlation for relationships using direct self-assessment measures was .30 (s^2 .0344), whereas for indirect measures it was .21 (s^2 .0204), excluding the outlier study. This difference is in line with the behavioural consistency approach (Wernimont & Campbell, 1968) where samples of behaviour are better predictors of criterion performance than signs or indirect measures. Of the studies reviewed, indirect measures were most commonly reported in the 1960's. Later studies generally used direct measures of the criterion behaviour.

The reviewed studies could also be classified into concurrent and predictive designs. Some investigations related self-assessment to criterion information collected concurrently (Henbest & Fehrsen, 1985) while others used self-assessment in a predictive role where criteria were more distal measures of performance or ability (Ekpo-Ufot, 1976). Such approaches to validation cannot necessarily be assumed equivalent for a number of reasons (see Guion & Cranny, 1982). However, it has been argued that the conceptual distinctions between concurrent and predictive designs have been exaggerated (Barrett, Phillips, & Alexander, 1981) and obtained validity coefficients are virtually the same (Schmitt, Gooding, Noe, & Kirsch, 1984).

The population correlation for the concurrent studies was .18 (s^2 .0333), rising to .29 (s^2 .0353), when the outlier study was omitted. The six predictive studies gave a population correlation of .28 (s^2 .0149). While there is some difference, the effect of the outlier study does not allow any clear conclusion regarding the strength of predictive versus concurrent studies.

Only three studies which used self-assessment in a predictive role in selection situations were located (Dicken, 1969; Mayo & Manning, 1961; Mischel, 1965). The population correlation was .19 (s^2 .0191), somewhat higher than other reviewers

suggest (e.g., Hunter & Hunter, 1984; Reilly & Chao, 1982). Each of these investigations used indirect measures. This population correlation did not include Farley and Mayfield's (1976) study as comparable statistical data were not available. Their study reported a nonsignificant relationship between a self-rating and earned commission using a sample of 1119 life insurance sales applicants (Table 4.1). No definitive conclusions can be drawn on the basis of these investigations; clearly further research is required in this area.

A further 11 studies were conducted in selection situations but involved relating self-assessment to other selection predictors not subsequent work criteria. The population correlation for these studies was .34 ($s^2 .0397$), the highest association detected in this review for any sub-group. Whether such a strong relationship would persist in similar predictive studies using more distal criteria remains a question largely unanswered by available research.

Factors Affecting Self-Assessment in Work Situations

A number of factors have been suggested as affecting self-assessment such as leniency or self-enhancement, halo, discriminant validity, rater perspective, and various moderator effects. This section will discuss these factors with particular emphasis on self-assessment in work and work-related situations.

We tend to overestimate many aspects of ourselves, for example height (Buckle, 1985) and driving ability (McCormick, Walkley, & Green, 1986). Self-assessments tend to be more lenient than assessments made by others such as supervisors, trainers, observers, and peers (e.g., Parker, Taylor, Barrett, & Martens, 1959; Shore & Thornton, 1986; Wheeler & Knoop, 1982). In work situations there is often a strong motivation to create a favourable impression (Davids & Pildner, 1958). In most cases, the papers

reviewed above showed that individuals rated themselves higher than they are rated by comparison groups (e.g., Anderson et al., 1984; Kirchner, 1965; Thornton, 1968). Some studies found no leniency effects (Levine, Flory, & Ash, 1977; Shapiro & Dessler, 1985) and Heneman (1974) reported self-ratings that were significantly lower than supervisory ratings on some dimensions.

Inflation of self-ratings seems to hold for different organisational levels (e.g., Holzbach, 1978; Shore & Thornton, 1986), for both trait and behavioural ratings (Kirchner, 1965), and for job related and person related dimensions (Parker et al. 1959). Leniency has been shown to be stronger for dissatisfied older employees (Lawler, 1967) and for low performers with high self-esteem (Baird, 1977).

Employee groups considered least promotable or ineffective are more likely to overrate themselves than more promotable or effective groups (Grande, 1966; Thornton, 1968). Primoff (1980) reported a study where operators' self-assessment correlated .61 ($p < .01$) with examination ratings during selection but the applicant with the highest self-rating actually failed the examination. Mabe and West's (1982) review, which included work and non-work related studies, found that of the 21 studies reporting information relevant to self-enhancement, 15 detected overestimation, four found no overestimation, and three found underestimation effects.

The level of common variance shared by several scales rated by the same person allows an evaluation of halo effects in judgmental ratings. In general, the reviewed studies showed that halo effects are less likely in self ratings (e.g., Kirchner, 1965; Wheeler & Knoop, 1982), but some studies have shown halo effects to affect both supervisor and self-ratings (Baird, 1977; McEnery & McEnery, 1987; Steel & Ovalle, 1984). Heneman (1974) suggests that although self-ratings show relatively small

amounts of halo-effect, there is little evidence to suggest that they fulfil the more severe requirements of discriminant validity. Small sample sizes, incomplete analyses, and the difficulty of meeting all three of Campbell and Fiske's (1959) requirements, as well as the relative lack of discrimination in ratings in general (Thornton, 1980), make it difficult to draw any definitive conclusions.

The discrepancies demonstrated between self and supervisory ratings (e.g., Brief et al., 1977) may be due to different aspects of performance being emphasised by the different types of rater. An examination of halo effects suggests individuals appear more able to distinguish between different skills, abilities, and knowledge than other raters. For instance, in a study of training needs assessment, factor analysis showed self-raters used three dimensions of performance whereas supervisors used only one (McEnery & McEnery, 1987).

A frequent outcome of dimensional analyses of self, supervisory, and peer ratings is that most of the variance is due to rating source rather than homogeneous groups of items (Holzbach, 1978). In addition, the level of agreement between self and supervisor ratings varies across different dimensions of work (e.g., Borman, 1978). Mount (1984) suggests that some dimensions are more reliably rated by one source than another as self or superior raters have different opportunities to observe performance, as well as different levels of expertise and experience. Christensen (1974) proposed that the dimensions rated most accurately are those a rater can understand and portray most vividly.

Few studies reviewed above considered gender as a moderator in respect of self-assessment. Where sex differences in self-perceived competence have been investigated, it has been shown that women are no less likely than men to have

positive self-referent, job-related attitudes (Snyder & Bruning, 1979). Some reviews have suggested that women have less positive self-referent attitudes (Maccoby & Jacklin, 1974) but these conclusions were not based on studies of employed persons in actual work settings. Dicken (1969) reported an isolated incident where females' scores on the Ego Strength Scale were more strongly related to field performance ratings than males. Nevertheless, in general, gender does not appear to affect the relationship between self and supervisory ratings (Ekpo-Ufot, 1979; Shore & Thornton, 1986).

Self-esteem appears to have some effect on self-ratings. Grande (1966) found that those rated as ineffective had higher self-esteem scores than those rated superior. King and Manaster (1977) showed that self-esteem was positively related to job interview performance with subjects selectively perceiving their performance in a way that was consistent with their level of self-esteem. Baird (1977) noted that subjects with a positive self-orientation were least likely to accept negative evaluations which did not fit with their own evaluations. His results showed that high self-esteem individuals who were rated as poor performers reported low satisfaction with their supervision. Tharenou and Harker (1984) also found global self-esteem was positively correlated with self-rated performance.

Social desirability is defined by Crowne and Marlowe (1964) as a need for social approval and acceptance and a belief that it can be obtained by culturally acceptable and appropriate behaviours. There is little research investigating social desirability as a possible moderator, which was surprising as one of the reasons suggested for practitioners being reluctant to use self-report information is their potential susceptibility to social desirability response bias (Arnold & Feldman, 1981). The limited evidence available, however, does not agree. In an experimental study using

work related ability measures, there was no evidence of social desirability moderating the relationship between self-assessment and ability test scores (DeNisi & Shaw, 1977).

The use of self-assessment as a predictor does not appear to affect minority groups adversely (Reilly & Chao, 1982). The differences between groups involved in the above review were generally nonsignificant (e.g., Ash, Levine, & Edgell, 1979; Brief, Aldag, & Sell, 1977; Levine, Flory, & Ash, 1977). Ash (1980) reported significant differences between racial subgroups on direct self-assessments of different typing skills. In general terms, self-assessments of typing abilities were lower for the minority group. In addition, minority self-assessments were less predictive of typing ability using alphabetical material but not significantly different from the majority group when self-assessing typing ability using numbers or tables.

Self-assessment may have some use in the identification and reduction of adverse impact. Primoff (1980) reported a study where the correlations between self-assessment and situational test items were calculated separately for black, Hispanic and white police applicants. Where there was a significant discrepancy between racial subgroups' item correlations, the item was withdrawn as a source of selection information. The rationale was that these items set up a different problem for different race groups so responses were not comparable.

Demographic characteristics other than gender and race as potential moderators of the accuracy of self-assessment receive little attention in work related studies. Research on educational achievement shows that more intelligent subjects tend to evaluate their abilities more accurately than less intelligent subjects (e.g., Bailey & Bailey, 1974; Kooker, 1974). Work related studies show no such moderator effects. Ekpo-Ufot (1976) reported educational level was positively related to ability ratings but no data

regarding moderating effects were given. DeNisi and Shaw (1977) found no moderator effects for general intelligence and Brief et al., (1978), including age and education among other demographic variables, found no significant effects.

Approaches to Improving Self-Assessment Validity

A number of methods designed to increase the validity of self-provided information have been suggested. According to Mabe and West (1982), one approach is to incorporate certain conditions of measurement into a self-assessment request. For instance, instructions could be included in self-assessment requests that would lead assessors to expect their self-assessments to be "validated" or compared to criterion data. Such "expectations of validation" may enhance the desire to be accurate rather than self-enhancing or lenient. Shrauger and Osberg (1981) have suggested self-assessment can be facilitated by specifying the exact behaviour and situations involved, facilitating recall of previous experience, and encouraging self-assessors to be accurate.

Some researchers have attempted to detect those individuals who are most likely to overestimate and correct such overestimates for degrees of leniency or inflation. Pannone (1984) grouped electrician applicants into "fakers" and "non-fakers" based on their self-rating of previous work experience on a nonexistent piece of equipment. Fakers had higher average self-ratings of experience and a lower correlation (.26, $p < .05$) with a selection test than non-fakers (.55, $p < .001$). Faking on such an item appears to have a strong relationship with lenient assessments and self-assessment validity. Anderson et al. (1984) used a similar approach but with a number of bogus items which were used to develop corrections for inflated self-estimates of work experience/ability. The corrections for inflation greatly improved the validity of self-estimates when compared with selection examination scores.

Other approaches to increase accuracy include focusing appraisal on performance feedback (Steel & Ovalle, 1984), avoiding direct self-assessment by asking for preferences on various job elements for matching with jobs (Cleff & Hecht, 1971), and requesting relevant autobiographical information which is then judged by others (Hough, 1984).

Steel and Ovalle (1984) compared self-appraisals based on an incumbent's view of him or herself and self-appraisal based on the performance feedback the incumbent received from a supervisor (FBSA). FBSA ratings displayed significantly less leniency, more variance, and more agreement with superior ratings than normal self-appraisals. A second study showed that the accuracy of self-appraisal increased in proportion with the amount of feedback given (Steel & Ovalle, 1984). Encouraging individuals to attend to information that describes what others think of their performance appears similar to encouraging self-focusing. Self-focused individuals are more likely to demonstrate consistency between self-reports and actual behaviour (Pryor, Gibbons, Wicklund, Fazio, & Hood, 1977).

The matching approach investigated by Ash, et al. (1979) asked candidates to choose which three of nine job conditions (e.g., standing, walking, and bending) they liked most and which three they liked least. Candidates were also asked to choose the six most liked and six least liked of 19 job tasks (e.g., filing). These choices were then compared to job condition and job task requirements to match candidates and jobs. Little adverse impact was demonstrated and an evaluation of the technique showed significant but low positive correlations of a match index with supervisory ratings (Levine, Edgell, & Ash, 1978). Job/person matching approaches, such as the Cleff Job Matching System (CJMS) which uses a self-interview checklist, has been described as

having predictive validity beyond that of "normal" psychological test batteries (Cleff & Hecht, 1971). The CJMS also has been established as a useful vocational guidance tool and a valid indicator of retention (Nathanson, 1975).

Asking applicants to describe major achievements which illustrate their knowledge, skill, and ability on behavioural and job related dimensions allows selective presentation of positive information. The accomplishment method of requesting and evaluating such self-reported protocols for attorneys has demonstrated relationships of .25 ($p < .05$) with job performance (Hough, 1984). This approach also performed better as a predictor when compared to inventories requiring judgments of action in job-related situations and judgment of relative importance of tasks (Hough, Keyes, & Dunnette, 1983).

Future autobiographies rated by judges have demonstrated relationships of .41 ($p < .01$) with evaluations of effectiveness for Peace Corp volunteers (Smith, 1966). Although research evidence is sparse, the validity of this approach appears to be affected by how the information is requested (Ash & Levine, 1985) and the dimensions used by judges for ratings (Tullar & Barrett, 1976). In addition, one study revealed an unusual moderating effect, religion. Ezekiel (1968) reported predictive validities were far stronger for Protestants than for Catholics.

This chapter has shown that researchers have examined self-assessment in a variety of settings and formats using a range of criteria. Reviews and individual studies reported in the literature show that self-assessment has predictive potential, particularly in the prediction of educational achievement and vocational choice. A number of attempts have been made to enhance the validity of self-assessment. The present review of research using self-assessment in work and work-related areas,

revealed only a sprinkling of studies scattered across a range of settings. There appears little in the way of a concentrated effort in specific areas, especially in the use of self-assessment as a predictor in personnel selection. Chapter Five describes a research project that is aimed at remedying the lack of empirical work in this field.

CHAPTER FIVE

AN EVALUATION OF SELF-ASSESSMENT IN PERSONNEL SELECTION

An examination of self-assessment in the context of personnel selection and from the perspective of the psychology of self in Chapters Two, Three, and Four has highlighted a number of factors. Firstly, individuals have extensive knowledge of their own abilities gathered in a wide variety of situations over a lengthy period of time. This knowledge forms the basis of self-evaluative judgments that have demonstrated acceptable predictive validity in a range of applications (e.g., Mabe & West, 1982).

Secondly, researchers have argued that self-assessment is a complex process, susceptible to distortion and selective justification (e.g., Snyder & Skrypnik, 1981). Self-assessments are sometimes required in situations where negative evaluations have a large potential cost for the assessor (Wicklund & Gollwitzer, 1983). Appropriate procedures and conditions, however, can reduce the effects of such biases, substantially enhancing the validity of self-assessment (e.g., Anderson, Warner, & Spencer, 1984; Mabe & West, 1982; Pannone, 1984).

Thirdly, research evaluating the validity of self-assessment in work and work-related situations is scanty, especially in the case of personnel selection applications (e.g., Shrauger & Osberg, 1981). Robertson and Makin (1986) feel that while personnel selection is a bilateral rather than a unilateral decision making process, it is curious that the participant with the most to lose from a bad decision is not given a more central role. As discussed in Chapter Four, reported research often does not directly request self-evaluations of performance, frequently compares self-assessment with

other selection predictors rather than more distal criteria, and is frequently conducted in experimental situations that have little in common with the realities of real life selection. The paucity of research examining the utility of self-assessment as a predictor in real life personnel selection situations, coupled with the predictive potential of such an approach, indicates a particular need for research in this area. As with any predictor being evaluated for use in personnel selection, however, validity is not the only issue.

The Acceptability of Self-Assessment in Institutional Personnel Selection

If a predictor is to be useful in a personnel selection situation it has to be both accurate and acceptable. It seems that in real life, acceptability is more important to practitioners than validity, especially when the popularity of a technique is used as a bench mark. Unstandardised interviews, for example, are highly acceptable but could hardly be described as accurate (e.g., Schmitt, 1976).

It has been proposed that self-assessment is rarely used in personnel selection because it is unacceptable. Levine, Flory, and Ash (1977) feel that this is because practitioners regard applicants as incapable of evaluating themselves and have a vested interest in a selection decision which would use self-assessment information. There is little empirical research supporting this judgment however. Rarity of use may be due to practitioners being unaware of such an approach or that they are unable to deviate from standard institutional procedures. Reports of investigations that do exist give little detail regarding any difficulties researchers encountered in having self-assessment accepted in the particular settings used.

An evaluation of predictor acceptability should be based on practitioners' responses to the introduction, implementation, and continued use of the technique in question. Such an approach would provide an estimate of acceptability in the work contexts in which the predictor could be used. A sample of practitioners could be asked to allow the implementation of self-assessment as part of their selection procedures. The proportions of those who firstly, agreed to implementation and secondly, who supported its continued use could then be calculated to provide an estimate of acceptability.

It may be that certain characteristics of the culture in which such an investigation was conducted could have an effect on how a new selection technique was received. Work is not something that exists in a vacuum but in the context of society. The effect of culture is reflected in the differences and similarities between values of workers in different countries (Hofstede, 1984). While difficulties exist in defining what culture is (Negandhi, 1983), its impact on work cannot be ignored.

The New Zealand Context

New Zealand as a small, somewhat isolated, multi-cultural society displays values reflected in attitudes to work which are different to those encountered in many Western societies (Ritchie, 1973). The New Zealand labour force of around 1.6 million individuals is about 42% of the total population (McLennan, Inkson, Dakin, Dewe, & Elkin, 1987). Approximately 1.3 million of the labour force describe themselves as full-time employees (Department of Statistics 1981). The country is many thousands of miles from the European, American, and Asian countries which are the major customers for its exports.

As described by Graves and Graves (1977), the New Zealand culture is made up largely of three broad ethnic groups: European, New Zealand Educated Polynesians, and Pacific Islander Immigrants. Europeans (pakehas) are New Zealanders from a British cultural background. New Zealand Educated Polynesians are in the main Maoris but also include those from different island nations who were either born here or who came as school children. Pacific Island Immigrants are those workers from various Pacific Island communities who have come to New Zealand as adults. Samoans, Cook Islanders, Tongans, Nuieans, and Tokelau Islanders are among those included in this group.

Feldman, Ah Sam, McDonald, and Bechtel (1980) have described the main differences that exist between representative ethnic groups as being based on the exchange of rewards, extended family obligations, the importance of individual wealth and property, and work organisation. The main differences are between European and Polynesian groups (Maoris, Samoans, Tongans, etc.).

1. Samoans and Maoris are more likely to exchange rewards on a personal basis (love, service, or status) whereas Europeans deal in more universal mediums of exchange (money, goods, or information).
2. The obligations Europeans feel toward their immediate family are felt by Polynesians toward their extended family which would include grandparents, great-aunts, great-uncles, aunts, uncles, in fact any person who could be regarded as members of one's household. Similarly, individual achievement brings honour to the family as well as to the individual.
3. Compared to Europeans, Polynesians are relatively unconcerned with personal wealth and property and are expected to share their possessions among others,

especially relatives. Social standing in a community is often influenced by an individual's generosity.

4. Polynesian work organisation is based around cultural lines of authority with those of high rank supervising others. In a European system authority is based largely on competence whereas in Polynesian systems social standing is also very important.

Many cultural differences are reflected in work situations (e.g., Barnes & Jamieson, 1977; Inkson, 1987; Nedd & Marsh, 1979, 1980) but not necessarily in all attitudes and behaviour reflected in work related measures. For example, both preference and liking of various work outcomes (Feldman et al., 1980) and worker preference for leadership style (Anderson, 1983) have shown few cultural differences.

In recent times a number of changes have made a substantial impact on the social and work scene in New Zealand. In the 1970s, working life was described by Hines (1976) as rather unique due to factors such as full employment, small companies, egalitarianism, the social welfare system, and the tax structure. Unemployment was maintained at well below 1% of the work force. The average company size was around 20 employees and there was no formal or recognised class structure. Medical and hospital services were provided for under a free health care system and the tax structure acted as a barrier to wealth accumulation. New Zealanders enjoyed a cradle to grave welfare system and a wide range of opportunities to choose what they did with their working lives.

According to a number of researchers, the effect on work attitudes was not necessarily positive. New Zealanders were described as lazy, passive, and unimaginative (Twinn, 1977); low in achievement motivation (Hines, 1974a); casual,

affiliation-oriented, and informal (Hines, 1976). In summarising the findings from a number of studies of job satisfaction conducted in the 1970s, Inkson (1987) suggested that many New Zealand workers had a rather passive, instrumental orientation to their employment. In a sense, their work existed to finance their leisure and provided an opportunity to relax after hectic weekends of activity.

McLennan et al. (1987) show that New Zealand of the 1980s has seen some major changes from the situation described above. Unemployment has become a fact of life with the demands of industry changing to needing fewer unskilled individuals and more technically competent, better educated employees. The Labour government of the mid 1980s has turned to a market orientated approach, reducing the level of state intervention in many facets of the economy. Various state controlled services previously part of the social welfare system have been opened up to private enterprise. This has resulted in competition for areas such as health which at one time was the sole preserve of the state. The deregulation of industry, the dismantling of state control, and the freeing of currency have given business in New Zealand a new confidence (Inkson, Henshall, Marsh, & Ellis, 1986).

It is somewhat unclear as to whether renewed business confidence will influence organisational attitudes to various practices such as personnel selection. The new market orientated approach has encouraged businesses to become more competitive and concerned with their profitability but it is unlikely that organisations evaluate any personnel selection procedures in cost - benefit terms. No organisations responding to Henderson's (1987) survey used any form of utility analysis. New Zealand practitioners are probably no different to their foreign counterparts who generally ignore the new methods of calculating financial costs and benefits of selection techniques (Cascio, 1987; Makin & Robertson, 1986).

Other factors which may influence the acceptability of self-assessment include the level of involvement psychologists have in industry, an absence of effective legal requirements for valid selection methods, and the multi-cultural nature of New Zealand society. Firstly, in comparison with other Western countries, few New Zealand organisations (7-16%) use psychologists (Bull & Spicer, 1983; Hines, 1972). Although a variety of reasons for such a low level of involvement in business has been suggested (e.g., Inkson, 1984), few New Zealand personnel practitioners seem to have any idea of what psychologists can offer. A lack of professional contact may make the introduction of any new approaches to personnel selection from a psychologist difficult although not necessarily hopeless. Both business managers and occupational psychologists regard personnel selection as an area of increasing importance for the future (Bull & Spicer, 1983; Mirfin & Hesketh, 1985).

Secondly, the legal sanctions against invalid selection procedures do not appear to be effective in New Zealand. General employment methods seem to ignore all but the most blatant discriminatory practices (George & Smith, 1986b; Smith & George, 1986a). Practitioners are under little pressure to change their current selection procedures so may not be interested in alternative methods. They appear satisfied with the interview supported by references and application forms, techniques which give little opportunity for the candidate to obtain job information and contribute to selection decisions.

Thirdly, the cultural composition of New Zealand may also affect practitioner and applicant attitudes to alternative selection methods. Spoonley (1977), for example, showed that many employers had overly negative views of Pacific Islanders as potential employees. Islanders were described as having no idea of work quality,

being reluctant to mix in the work situation, having hygiene problems, and a poor understanding of English. Spoonley's study also showed that Europeans were given access to jobs in preference to Pacific Islanders who were similarly or even better qualified. It seems unlikely that employers reporting and demonstrating such attitudes would wish to place any credibility in self-evaluative information supplied by Islander applicants.

An interesting but rare possibility is that for some New Zealand organisations, selection decisions will be made by the applicants. Increasing unemployment has been a particular problem in the rural areas of New Zealand and one approach to increasing commercial activity and employment has been to form cooperative businesses. The cooperative approach reflects Maori tradition and the system under which many earlier pakeha settlements were run (Hutchinson, 1984). In these participative businesses individual workers decide whether or not they want to be "employed". A recent case described by Pirikahu, MacPherson, Gibbs, Kahu, and Ponter (1984) illustrates the nature of a New Zealand cooperative.

The town of Patea lost its one major local industry, meat processing, resulting in high levels of unemployment and reducing the cash flow into local industry and retail sectors. To increase commercial activity the Patea community, along with the local Meat Workers Union, formed the Rangitaawhi Marae Enterprise Trust, an organisation based on the "Turangawaewae" (culture and traditions) of the "Tangata whenua" (the people of standing). The Trust leaders wished to revive the economic base of the land of their "tipuna" (ancestors) and not be subjected to an outside authority. They were committed to courtesy and equitable treatment of all and had a solidarity based on their experience at the freezing works, "being literally an extended family" (Pirikahu et al., 1984; p. 53). On this basis the trust developed a number of

enterprises involving horticulture and rabbit farming. The activities of the Trust and other related organisations, such as the Patea Maori Club, has given a purpose and dignity to its previously unemployed members. Trust members have taken control of their working lives.

An important part of the evaluation of an alternative selection technique is an assessment of its acceptability in the context for which it is intended. Apart from Spoonley's (1977) work on employers' attitudes toward different cultural groups, information on the attitudes of New Zealand personnel selection practitioners appears largely ignored by researchers. As discussed in Chapter One, there appears little variation in selection methods beyond the interview supported by references and an application form. Whether this suggests a resistance to alternative selection techniques is difficult to say. In the absence of such information it is important to include an assessment of the acceptability of self-assessment as part of an evaluation of its utility for New Zealand institutional selection.

Aims of the Research Project

The present investigation is aimed at evaluating and extending an approach demonstrated in laboratory and other settings. It is known that self-assessment is a valid predictor under certain conditions but what available research seems unable to show is whether or not self-assessment has predictive utility in personnel selection. In addition, there is little information on the acceptability of such an alternative selection procedure in New Zealand. The purpose of this investigation is to evaluate self-assessment as a personnel selection predictor in New Zealand institutions. Specifically this will involve the following:

1. An evaluation of the acceptability of self-assessment as a selection tool to New Zealand practitioners.
2. An evaluation of the unique predictive variance offered by self-assessment above that provided by selection methods used in New Zealand organisations.
3. An evaluation of possible moderator effects on the self-assessment - work criteria relationships. Potential moderators include cognitive ability, gender, educational level, age, social desirability, locus of control, level of job information, self-assessment experience, and whether or not self-assessment requests are made in "compared to others" or absolute terms.

The Acceptability of Self-Assessment in Personnel Selection

To evaluate the acceptability of self-assessment in New Zealand, a range of organisations were approached and asked if they would be willing to include self-assessment in their current personnel selection procedures. A number of factors influenced the type of organisation approached. Firstly, it was intended that the sample include representatives of each of the major industry groups to ensure employers in each of the major sectors were represented. The classification adopted was that used by the Department of Statistics (1981) in reporting census information. Secondly, it was also intended that the selection procedures of those companies agreeing to participate in the research project would include a range of techniques commonly used in New Zealand. Personnel selection in this country generally involves an interview supported by an application form and references or referees' recommendations. In some instances work samples and pencil and paper tests are used (Chapter Two).

Thirdly, it was felt that larger employers were more appropriate for the purposes of this research project. Organisations employing 250 or more employees were expected to have a centralised personnel function, a manager with designated responsibilities and standardised selection procedures as well as an ongoing recruitment and selection programme (New Zealand Institute of Personnel Management, 1979). A lack of standardised procedures or centralised control may have resulted in a variety of selection procedures within an organisation with no standard selection method for comparison with self-assessment.

Method

Settings

Suitable organisations were identified using New Zealand business directories (e.g., Fourth Estate, 1984) and personal contact with researchers and practitioners involved in personnel selection. Sixty nine organisations were approached by letter asking for cooperation in the research project. The organisations approached represented virtually all of the major industry groups (see Table 5.1). Fourteen companies had more than one major function and therefore contributed to more than one industry group.

Table 5.1

Sample of Organisations Classified Into Major Industry Groups

Industry Grouping	Sample Proportion	Overall NZ Proportion
Agriculture, Hunting, Forestry, & Fishing	9.7	10.8
Mining & Quarrying	2.4	0.004
Manufacturing	48.2	23.5
Electricity, Gas, & Water	1.2	1.1
Building & Construction	2.4	6.4
Wholesale, Retail Trade, Restaurants, & Hotels	14.5	16.5
Transport, Storage, & Communication	4.8	8.1
Financing, Insurance, Real Estate, & Business Services	12.0	6.9
Community, Social, & Personal Services	4.8	23.1
Activities Not Adequately Defined	0.0	3.6

Procedure

Each organisation was approached by letter requesting cooperation in a research project evaluating the contribution of self-assessment to personnel selection (see Appendix 1). Those organisations who were willing to participate were visited and the project was explained in detail, including the nature of the self-assessment request and administration. Requirements of the research project were described including access to information on selection procedures and decisions, job performance criteria (including absenteeism and turnover), and expected levels of recruitment over the

following two years. The benefits to the organisation were outlined including an evaluation of present selection procedures and recommendations for possible improvements.

If an organisation agreed to the implementation of self-assessment as part of their selection procedure, a request form and administration procedure was developed for and approved by each organisation. The form for each organisation was printed and personnel staff involved in issuing the self-assessment request to applicants were instructed in its use. The researcher kept regular contact with organisational staff taking responsibility for the research project over the period for which the project ran.

The point at which criterion data were gathered depended very much on individual organisations. It was intended that self-assessment requests remain part of organisational selection procedures for a two year period in each setting.

Organisational measures of work performance were collected as they became available and data collection for each setting finished approximately three years after the first self-assessment requests were made. In addition to collecting criterion data, individual characteristics data (e.g., age, gender, education, work experience, etc.) were gathered as access to organisation files allowed.

Results

Of the 69 settings approached for cooperation in the research project, 21 (30.4%) declined the request. The reasons cited were, company policy, restructuring, low expected turnover (two to six employees per year), a lack of staff resources, a previous uncompleted university research project, and in two cases an expectation that

self-assessment was of little or no use. A total of 24 (34.8%) companies did not reply to the initial enquiry. A further 24 (34.8%) expressed an interest in the research and meetings were held with management as described earlier.

As a result of the initial meetings it was mutually agreed in eight settings that the circumstances were not suited to this research. The main reason for not continuing the project was a low level of recruitment (less than five per year) due to the reduction of staff through the use of attrition and minimal recruitment. The 14 remaining organisations agreed to participate in the research and in all cases implemented the self-assessment as part of their selection procedures. This involved 16 separate samples with some settings involving more than one employee sample (Table 5.2).

The range of selection methods used by the organisations agreeing to participate in the research appeared to adequately represent selection techniques commonly used in New Zealand (see Table 5.2). Interestingly, one setting, the Polytechnic, did not use an interview for selection. In addition, the Food Processing Company used an "availability check" which could be described as an extremely brief interview aimed at establishing when an applicant would be available to start work.

All settings used application forms, 14 used an interview format (87.5%), two used a work sample (12.5%), and one used a psychological test. Although only one setting used a standard reference form, all other settings required applicants to supply names of individuals to be used as referees. In the case of the Finance Company (No. 3) and the Food Processing Company (No. 8) the selection procedures used for the two main groups involved were very different; so these groups were considered separately.

ble 5.2

Organisations in Which Self-Assessment Research was Conducted

Organisation Product/Operation	Applicants	Current Selection Procedures	Sample Size Expected
Polytechnic	Nursing Trainees	Application Form, Reference Form	255
Petroleum	Managerial Staff	Application Form, Interviews	60
Bank	Clerical Staff	Application Form, Interviews	100
Bank	Managerial Staff	Application Form, Interviews	60
Metal Fabrication	Manufacturing Staff	Application Form, Interviews	75
Finance/Banking	Managerial Staff	Application Form, Interviews	55
Electronics Manufacturing	Assembly Line Staff	Application Form, Interviews Psychological Test	65
Credit Card	Service Staff	Application Form, Interviews	85
Food Processing	Seasonal Workers	Application Form, Availability Check	700
Food Processing	Managerial Staff	Application Form, Interviews	50
Shipping Company	Middle Management	Application Form, Interviews	40
General Retailing	Service Staff	Application Form, Interviews	50
Motor Vehicle Manuf'	Middle Management	Application Form, Interviews	50
Petroleum	Management Staff	Application Form, Interviews	50
Govt. Department	Advisory Officers	Application Form, Work Sample, Panel Interviews	80
Home Appliance Manufacture	Assemblers Workers	Application Form, Work Sample Tests, Interview	65

During the "life" of the project a number of organisations were dropped from the study. The reasons for withdrawal are listed below. Each is numbered as shown above in Table 5.2.

Petroleum Company (No. 2): Part way through the first year of the study this company transferred its personnel manager overseas and then through transfer or resignation lost its remaining personnel officers. A major reorganisation of the personnel function resulted and on request from the company the research was suspended until new personnel staff were installed. One year after this suspension the company stated that it would be unable to continue data collection in the foreseeable future.

Finance Company (No. 3): Two studies were implemented in this setting. Prior to final printing of self-assessment request forms the company suffered large (and apparently unexpected) financial losses. As a result a number of major changes in the organisation structure were made, aimed at increasing the accountability of divisional managers. These changes involved removal of the centralised personnel function. The personnel manager in charge at the time of the change stated that it would no longer be possible to continue the project as the company did not want to increase further the responsibilities of their divisional managers.

Metal Fabrication (No. 4): After the project had been running for 10 months and three completed self-assessment forms had been received, the personnel manager stated there would be little if any employment of staff in the next six to eight months. At the same time the manager reported that three applicants were unwilling to complete the self-assessment request as they felt unable to assess their performance on a job they had not done before. They were persuaded to complete the self-assessment. No further forms were received and as contact could not be

made with the personnel manager or his associates, the organisation was withdrawn from the project after 18 months.

Finance Company (No. 5): Two months after implementation the personnel function was reorganised, apparently due to the personnel manager resigning. The personnel officer directly responsible for selection reported that self-assessment was being used but that the organisation had only employed "one or two new staff". Despite repeated requests the officer was not able to attend meetings arranged with the researcher or pass on any completed self-assessment requests. Twelve months after the study had been in progress the personnel officer involved resigned. The replacement officer informed the researcher that no applicants had been given the self-assessment request and that the organisation did not want to continue the project as he saw little value in the approach.

Credit Card Company (No. 7): Immediately following development of self-assessment requests and procedures the personnel manager was promoted and a new manager employed. The new person was introduced to the project and agreed to continue the research after a three month "settling in" period. At a meeting four months after the project had begun, five completed self-assessment requests were received. However, despite continued telephone and personal contact no further requests were passed on. Ten months later the General Manager (Personnel) stated that the organisation no longer had the resources to continue the project. It appeared that this was due to the personnel manager's unwillingness to use the self-assessment requests.

Food Processing Company: Managerial Staff (No. 8): This sample involved 39 employed managerial staff at six different locations. The selection procedure was coordinated by one manager but performance appraisal was the responsibility of

managers at the separate operational sites. Throughout the study performance appraisals using a standard organisational format required by the company were difficult to obtain. Managers simply did not bother to complete the forms and in a number of cases managers and staff who were part of the sample resigned before any criterion data could be collected. Despite numerous attempts by the coordinating manager only 11 completed performance appraisal forms were obtained. This study was not analysed due to the small proportion of the sample for whom criterion data were available.

Shipping Company (No. 9): The sample expected was the lowest of all settings, but due to an apparent downturn in the organisation's business activities only two managers were employed in the first 12 months of the project. As no increase in employment was expected at that stage, it was agreed that the research be terminated.

General Retail Company (No. 10): During the first 12 months of the project 13 shop assistants were employed. At this stage the company sold all grocery retailing interests reducing staff by approximately 70%. As little, if any, recruitment was expected during the next 12 months because the company had to absorb a number of otherwise redundant staff, the organisation withdrew from the project. As criterion data were available for only six of the employed staff, no further analysis was done.

Motor Vehicle Manufacturer (No. 11): Four months after development and implementation of the self-assessment request the Supervisor for Recruitment and Salaries left the organisation. Her successor stated that the company wished to withdraw from the project as he did not feel it would be of any benefit to the organisation.

Petroleum Company (No. 12): Eight months after the self-assessment request had been developed and implemented, the employment officer requested a meeting with the researcher. He stated that there had been a number of negative reactions from applicants to middle management positions when they were asked to complete self-assessment requests. The self-assessment procedure in this organisation had involved giving applicants the request along with a stamped, addressed envelope for posting directly to the researcher. Applicants reacting negatively were concerned that the ratings they made would be passed back to the company despite the reassurances made by the employment officer and statements in the self-assessment request. The employment officer agreed to continue the project but on being contacted one month later said he had decided to withdraw. Eight completed requests were received by the researcher and as no criterion data were to be made available this setting was discarded.

Home Appliance Manufacturer (No. 14): Six months after the project had begun the company was taken over by a rival manufacturer. Two months after the takeover the Training Officer responsible for assembly line worker selection was made redundant. All but three completed self-assessment request forms, along with selection and training data, were destroyed before the researcher was able to gain access to the information. The Administrative Officer who assumed responsibility for selection stated that the new owners did not want to continue the project.

Overall, 12 (75%) of the original 16 settings where self-assessment requests were implemented did not produce suitable data which could be used to evaluate statistical relationships. Three of the remaining four samples were split into two as there were changes in either the selection procedure, the self-assessment request procedure, or

changes in both. The resulting analysis involved seven separate studies and 944 applicants as described in Table 5.3.

Table 5.3

Organisation Settings, Samples, Selection Procedures and Criteria Used in Analysis

Organisation Product/Operation	Applicants	Selection Method	Criteria	Sample Size
Electronics Manufacturing	Assembly Line Staff	Application Forms, Interviews, Psychological Tests	Induction, Performance Ap' Turnover	61
Polytechnic				
Study A	Nursing Trainees	Application Forms, Referee & Self- completed References	1st Year Success Course Success	129
Study B	Nursing Trainees	Application Forms, Referee & Self- completed References	Total Course Success	90
Food Processing				
Study A	Seasonal Workers	Application Form, Availability Check	Performance Ap' Absence, Tenure	176
Study B	Seasonal Workers	Application Form, Availability Check	Performance Ap' Absence, Tenure	375
Govt. Department				
Study A	Advisory Officers	Application Form, Panel Interview, Written Exercise	Promotion, Tenure	28
Study B	Advisory Officers	Application Form, Assessment Centre	Performance Ap' Promotion	85

Discussion

From an initial request to 69 organisations, 24 or 34.8% expressed an interest in the research. Two companies stated that it was their standard policy not to become involved in research projects, one in particular citing an instance where a researcher had failed to report back after that company had invested a considerable amount of time in a project. To what extent this response rate was due to self-assessment being unattractive as a selection tool is difficult to tell. Two organisations stated they did not have much confidence in such an approach so in these cases the effect was negative. On the other hand, during initial discussions with potential participants, all managers stated that they had agreed to a meeting because they were intrigued with the self-assessment idea.

In general, discussions with managers suggested that while they felt that applicants were able to make judgments about their own future performance, the assessments would be inflated. There also seemed to exist an expectation that the research would use some psychological subterfuge to avoid biased assessments. Whether this was from a popular conception of psychologists as people who used tests and other analytical devices to reveal otherwise hidden information about applicants or simply due to the research project taking a different approach is not known. Compared to some other countries, New Zealand organisations rarely use psychologists (e.g., Bull & Spicer, 1983) so it is quite possible many managers had little or no contact or knowledge of the activities of occupational psychologists.

Of the 16 settings participating in the project the reasons for withdrawal varied. In one case, company reorganisation removed a centralised personnel function leaving over 100 divisional and branch managers to select as they saw fit. Had these conditions existed two months earlier, the lack of both a standardised selection

procedure and centralised control of the personnel function would have meant the setting was not suitable for the research project. A severe downturn in recruitment was the major factor in two cases. Had the actual level of recruitment been known prior to developing self-assessment requests, these settings also would not have been part of the project.

In five cases reorganisation and resignation of staff placed new individuals in control of recruiting who did not want to continue the research. The replacement of the original company "contact" person who took managerial responsibility for the project was the most important factor in the attrition of research settings. In all cases where such a replacement occurred during the expected "life" of the research project, the organisation terminated its involvement. New managers when approached seemed more interested in settling into their new positions. It appeared that they did not have the time to be involved in a project that had been initiated by their predecessor and was in their eyes peripheral to their job. Two managers stated that they did not feel the project was of any benefit to their organisation. While organisational access is difficult to arrange (e.g., Chadwick-Jones, Nicholson, & Brown, 1982) it is apparently more difficult to maintain agreed arrangements, especially when personnel responsible for the research change.

Two settings terminated the research because applicants had been suspicious of the self-assessment request. The personnel administering the requests felt that such applicant reaction may prejudice their chances of employing certain professions whose skills were in short supply. In one setting (Table 5.2, No. 12) it was clearly shown on the request form that the procedure was part of a university research project rather than a company requirement. Applicants completing these forms were asked to post them on completion to the researcher in the stamped addressed envelope supplied.

Based on figures supplied by the company, approximately 50 requests were given to applicants and 39 completed forms were received by the researcher, a return rate of 78%. This may be a reasonable rate for mailed questionnaires but does not provide complete information for selection decisions. While some applicants may have simply forgot to post the requests, others may have been suspicious and held on to them until they knew the outcome of their application. The range of possible motivations for not returning the requests are many. The other organisation in this category (Table 5.2, No. 4) had the request presented as part of normal selection procedures. The manager making the requests reported that some applicants had asked "was it some sort of trick?" and although the applicants were persuaded to complete the forms they still seemed to be suspicious. This organisation was concerned about applicants' attitude to the requests and the researcher was unable to make contact with the personnel manager following this report.

It appears that the suspicion surrounding self-assessment on the part of practitioners may also exist in the minds of some applicants. A personnel officer in one organisation said that a couple of applicants felt such an assessment was the company's job not theirs. It seems that asking for a self-assessment may have violated what had been regarded as the norms of the "selection game".

Based on the response rate and the attrition observed in the research project, it would appear that more organisational settings are likely to reject rather than accept the use of self-assessment as a predictor in personnel selection. While the response rates and attrition observed can be attributed to other factors such as New Zealand organisations not wanting to be involved in research irrespective of subject matter, it seems that self-assessment does not have the same attraction for practitioners that other selection methods seem to have. In this respect, the judgments passed by other

researchers regarding the acceptability of this approach (e.g., Levine, Flory, & Ash, 1977) have, in part, been supported.

Interestingly, a byproduct of the self-assessment requests was commented on in two settings. Personnel administering the requests involving managerial staff (Table 5.1, Nos 2 & 8) stated that they felt the self-assessment request provided little useful information as far as selection was concerned but it was exceptionally useful in identifying an applicant's training needs. By discussing self-assessment responses with an applicant they were able to establish what type and amount of training would be required in what was described as a cooperative rather than confrontational atmosphere. In addition, applicants reportedly took responsibility for their own training with the support of organisational resources as a result of these discussions.

Seven separate samples remained in the research project for the intended length of time. It is these studies which allow an evaluation, in quantitative terms, of the validity and moderators of self-assessment. The results of these studies are reported in the next four chapters.

CHAPTER SIX

VALIDATION OF SELF-ASSESSMENT VALIDITY IN AN ELECTRONICS MANUFACTURING COMPANY

The electronics manufacturing company was part of an organisation with similar plants in Australia. The company employed approximately 350 staff engaged in the manufacture, design, and development of electronic equipment (e.g., computers, amplifiers, magnetic tape decks, radios, speakers, monitoring devices, etc.).

Method

Subjects

The subjects were 61 applicants who accepted an offer of a job as an assembler for an electronics assembly company from a total applicant pool of 153. Four of the total applicant pool were not able to complete the self-assessment request form and a further 88 were not offered jobs. The average subject age was 25 years and three months with a range of 28 years eight months (44 yrs 8 months - 16 yrs). Four cases (6.6%) were male, 57 (93.4%) were female. Few cases had educational qualifications at the secondary school level. One applicant had sixth Form Certificate and four had School Certificate. The sample was classified into three main racial groups, Pacific Islander, $n = 18$ (29.5%), New Zealand Maori, $n = 19$ (31.2%), and European or Pakeha, $n = 22$ (36.0%), with two applicants falling into an "Other" group (Cambodian and Indian).

Procedure

Development of self-assessment requests. An assembler was responsible for designated assembly tasks such as soldering or clipping electronic components in a

pre-set sequence on to a circuit board and assembling the major components of radios (e.g., chassis, fascia, control knobs, etc.) in a production line arrangement. All assembly work was closely supervised and checked by testing staff who were able to identify faults and trace them to the individual assemblers responsible. Apart from a brief verbal description of the job given to candidates during the selection procedure, no job content information was available to candidates. The design of the self-assessment form, which can be seen in Appendix 2, was based on the following points.

While ensuring applicants have sufficient access to job information that is normally available, it is important in this type of intervention that no additional information is provided by the self-assessment requests themselves. This is to ensure that the effects of the investigation on normal selection procedures and applicant decision making is minimised. Self-assessments are made on the basis of information available to the candidate, not the information ideally available to the organisation such as contained in a job analysis.

The conditions set by the company required that the self-assessment request be developed in conjunction with personnel staff to ensure its acceptability. In addition, the request was not to require undue additional staff involvement during selection procedures. Accordingly, the staff time required was kept to a minimum by designing the self-assessment request as a self-administered form. This format also standardised requests and required minimal training for personnel staff.

In line with Mabe and West's (1982) meta-analytic review, the self-assessment request was couched in relative terms, included information aimed at enhancing an applicant's expectation of validation, and was matched to work performance criteria. The comparative characteristic was incorporated by instructing applicants "when making

this rating compare yourself with the sort of people you think would be doing this type of work". To encourage expectations of validation the statement, "the rating you give yourself now will be compared to your performance on the job", was included in the request. The self-assessment request was matched to work performance criteria by asking applicants to directly assess their own future job performance.

The self-judgment of one's ability to perform a job is equivalent to the assessment made by the potential employer on which he or she bases a job offer decision. The focus of the present investigation was the overall judgment of future work performance, not a number of self-assessments on different dimensions of a job which are somehow combined to form a predictor. Although some researchers feel that the concept of a single criterion is unrealistic, as job performance is a function of a number of dimensions (e.g., Dunnette, 1963a), others would support an overall approach. Lawler (1967) suggests that when people are asked to make global ratings they are efficient processors of critical-incident data. Their simple global performance ratings appear to yield a reasonable approximation of what would be obtained by using a more extensive rating procedure.

Factor-analytic studies of ratings have invariably produced a single job performance factor (e.g., Whitlock, 1963). Supervisors and self-raters evaluating job performance tend not to differentiate between different dimensions of performance. One recent example involving a factor analysis of self and supervisory ratings of actual performance confirmed this (McEnery & McEnery, 1987). For these reasons self-assessments of global job performance were requested.

A draft self-assessment form was developed, based on the performance appraisal forms used to in assembler induction training and on-job appraisals made during the fourth

and seventh week of employment (see Appendix 3). The instructions in the self-assessment form told applicants that the request was part of a research project and was not used by the employer when making a selection decision. As well as the overall self-assessment of job performance, applicants were asked how well they would do various aspects of the job (Timekeeping, Social Acceptability, Social Interaction, Work Area Tidiness, Care of Hand Tools, Work Attitude, Care of Product, Work Speed, and Use of Tools) on a scale of 1 (Poor) to 5 (Good).

Two items used on appraisal forms (soldering and component recognition) were not used for self-assessment as personnel staff felt applicants would not have a very accurate idea of what these particular aspects of work involved. The self-assessment request was pilot tested on five current employees during working hours and ten applicants during normal selection procedures. No problems were reported or observed in completing the form apart from two applicants who could not read. As the company required employees to read and write in English, no changes were made to the format.

As well as the overall self-assessment of job performance (SAJOB), an average of self-assessments (SELFAV) made on the different job aspects provided an additional measure of self-assessment in this setting.

To give applicants the opportunity to gather as much information as the selection procedure provides, the self-assessment request was included as the last stage of the selection process. While one cannot state with absolute certainty that organisational selection decisions and applicant decisions are made at the point when all relevant information has been gathered (e.g., Springbett, 1958), placing the request at the

conclusion of a selection procedure means that the applicant's assessment is made at around the same stage as the organisational decision.

Criteria. The three criteria used were induction ratings, work performance ratings, and turnover.

Each new employee participated in a four day induction course dealing with the basic skills they would immediately require on the assembly line (e.g., care of product, work speed, use of tools, etc.). At the end of the course the training supervisor rated each participant on those skills taught. As the content of each induction course depended on assembly line requirements, the induction rating (INDUCT) was the average of these rated skills (Appendix 3).

Ratings of work performance were made after the fourth and seventh week of employment using the same form (Appendix 3). The ratings made at the seventh week was a confirmation or modification of the earlier rating, not a further independent assessment of performance. The average rating of skills (PERFORM) made during the seventh week was used as the work performance measure for those cases still employed.

For the turnover criterion (TURN), the sample of 61 was classified as either "Stayers" (tenure = one year or greater), "Voluntary Turnovers" (e.g., resignation, retirement) and "Non-voluntary Turnovers" (e.g., dismissal). This classification was based on the functional approach to classifying turnover (Hollenbeck & Williams, 1986; Muchinsky & Tuttle, 1979).

Moderator variables. "A variable is a moderator if the relationship between two (or more) other variables is a function of the level of the moderating variable" (James & Brett, 1984). The detection of moderators is of practical and theoretical importance in personnel selection (Ghiselli, 1972). As discussed earlier, a number of individual characteristics may moderate the relationship between self-assessment and a criterion measure. Those investigated in this study are outlined below.

a) Self-evaluation experience.

According to Mabe and West (1982), self evaluation is a skill that could be improved by practice. Applicants who have had the experience of self-assessment should be more accurate in their evaluations than those who have not had such an opportunity. Within the constraints of organisational selection procedures it was not possible to give applicants experience in self-assessment. As a result, the effect of self-assessment experience was evaluated by classifying applicants into two groups, those that stated they had self-assessment experience and those that did not.

b) Source of job information.

The accuracy of self-assessment may be affected by the type of information available to the rater. Bandura's (1977b, 1978, 1982a) self efficacy theory divides the sources of information on which assessment of one's performance is made into enactive experience, vicarious experience, verbal persuasion, and emotional arousal. As discussed earlier, the theory suggests that the quality of the information varies depending on the source resulting in an effect on the accuracy of self-assessment. For ethical and practical reasons it was not possible to manipulate the type of information available to applicants. To investigate the effect of differential

exposure to job information, applicants were asked to indicate what sort of information they had about the job. The response categories were:

- | | | |
|----|--|-------------|
| 1. | By working "in the job" myself | (enactive) |
| 2. | By watching others work "in the job" | (vicarious) |
| 3. | By being told by others about working at | (verbal) |
| 4. | Through any other source | (other) |

Emotional arousal data were not collected due to the difficulty of conveying the concept to applicants.

c) Social Desirability.

The 33 item Crowne and Marlowe (1964) Social Desirability Scale (SDS) was used as a measure of the degree to which applicants tended to describe themselves in socially desirable terms. This scale has been used extensively and its validity has been demonstrated in a number of experimental settings ((Robinson & Shaver, 1973). It is particularly suitable for a normal population in that it does not contain items referring to psychologically abnormal behaviour as in the case of the scale developed by Edwards (1957).

To investigate other possible moderator effects data were collected on age and race. Subjects aged 22 years eight months or older formed the upper age group and subjects aged 22 years six months formed the lower age group. Subjects scoring 23 or more on the SDS formed a "high score" group and subjects scoring 22 or less formed the low score group. Subjects were classified as either European (n = 22, 36.1%), Maori (n = 19, 31.2%), or Pacific Islander (n = 18, 29.5%). Two (3.2%) applicants fell outside these categories and were not included in this analysis. Predictor - criteria

relationships examined in the overall analysis were calculated for each moderator subgroup. Gender was not included as there were only four males in the sample. School qualifications as a variable also was not tested because only five of the sample had any school qualification.

Analysis. Regression analysis is a general multivariate technique which allows the analysis of the relationship between a dependent or criterion variable and a group of independent or predictor variables. It is statistically similar to Linear Discriminant Function Analysis (LDF) apart from the nature of the criterion variable. Regression is designed to predict a position on a criterion scale whereas LDF is used to classify cases into categories. As Harris (1985) puts it, multiple regression is nothing more than the familiar Pearson correlation between an outcome measure or dependent variable and a linear combination of a subject's scores on a number of predictor variables. A measure of the accuracy of prediction, or strength of linear association, is the ratio of explained variation in the dependent variable, Y , to the total variation in Y , or R^2 (Edwards, 1984). The contribution of predictor variables to a linear association described by regression coefficients is generally reported as BETA weights, or standardised regression coefficients. Standardisation allows a comparison of two or more independent variables when these variables are measured in different units (Kim & Kohout, 1975). The effect of an additional predictor variable being added to an equation can be described by the change in R^2 . R^2 is a part correlation coefficient being a description of the relationship between the dependent variable and the additional predictor, with the linear effects of the variables already in the equation removed (Norusis, 1985).

The object of LDF is to weight and linearly combine scores on a number of predictor variables so the criterion groups which form the dependent variable are as statistically

distinct as possible (Klecka, 1984). The discriminating power of a LDF function is described by the canonical correlation, a measure of the relationship between the function and group membership. In addition, the accuracy with which a sample of cases is classified into dependent variable groups is a further indication of the efficiency of the discriminant function (Klecka, 1975). Discriminant function coefficients can be used to describe weights calculated for each predictor variable. Standardised procedures give statistical descriptions of the discriminating power of the function, describing which independent variables are being used and what contribution they are making. Once a function has been formed it can be used to classify both the cases used in the analysis and new cases.

These multivariate techniques were considered appropriate tools for describing the data gathered in this study because of their recognised robustness (Harris, 1985; Klecka, 1984; Tabachnick & Fidell, 1983) and the extensive descriptive information they provide. Regression analysis¹ was used to examine the impact of self-assessment (SAJOB and SELFAV) on INDUCT and PERFORM in addition to the effect of the organisational selection decision (GRADE). As turnover was treated as a nonlinear categorical variable, linear discriminant function analysis (LDF) was used to evaluate the self-assessment - TURN relationships.

The way in which moderator effects were evaluated requires some discussion as there is little agreement as to how potential moderators should be analysed. An examination

¹ There appears no clear answer in the literature as to what sample size is sufficient for multi-variate analysis. Gorsuch (1983) recommends a minimum of five subjects per predictor variable. Harris (1985) recommends that where predictors and criterion have grossly non-normal distributions and the number of cases minus predictors is less than 50, such analyses should be interpreted with caution. In the light of this, and other statements, it was felt that the sample size used in this study was sufficient for multivariate analysis to provide descriptive data.

of the literature shows that the issues surrounding the statistical and analytical techniques used for evaluating moderators have not been resolved (e.g., Guion, 1976; Peters & Champoux, 1979; Zedeck, 1971).

There seems to be a growing trend toward the use of moderated regression analysis for moderator detection in larger samples (e.g., Arnold & Evans, 1979; Evans, 1985). Moderator analysis is essentially a test for interactions. As described by Cohen and Cohen (1983), "an interaction between two variable sets U and V is represented by multiplication of their respective independent variables and then linearly partialling out the U and V sets from the product set. A hierarchical analysis accomplishes this partialling procedure automatically, for example, entering U, V and then UV. The increment due to R^2 due to UV in such a sequence is thus attributable to the U by V interaction" (p. 349). This approach has been demonstrated as a reliable method of identifying interactions in samples that have sufficient numbers for multivariate analysis (Evans, 1985).

An equivalent detection procedure for smaller samples is not anywhere near as clear cut with approaches such as quadrant analysis (Hobert & Dunnette, 1967), differential validity (e.g., Fredriksen & Melville, 1954), and differential predictability (Ghiselli, 1956) available to the researcher. Each method has been both criticised and supported on methodological and practical grounds (e.g., Abrahams & Alf, 1972; Dunnette, 1972). Few examples of moderator analysis with small samples are to be found in the occupational psychology literature but where they do exist the differential predictability method is generally used (e.g., Brief, Aldag, & van Sell, 1977). This approach uses a third or moderator variable as the basis for splitting the total sample into subgroups. The degree of difference between the validity coefficients of the subgroups determines whether or not the third variable is acting as a moderator.

Problems do exist with the subgroup method, including the loss of statistical power when scores on a continuous moderator variable are collapsed into two or so categories (Finney, Mitchell, Cronkite, & Moos, 1984). There are also situations where subgroups may have a relatively larger range of predictor and criterion scores (than the total sample) so inflating the validity coefficient without increasing predictability (Abrahams & Alf, 1972).

Where a discontinuous moderator such as sex is used, a comparison of validity coefficients for a significant difference is appropriate (Arnold, 1982). Despite the problems with continuous variables, differential predictability appears the better choice for smaller samples as there is apparently no clear advantage offered by alternative methods. In addition, such an approach can be directly compared with the other occasional moderator analyses on similar sized samples in this field.

Moderator effects of the predictor - TURN relationships in this investigation were not evaluated because sub-group numbers were insufficient for LDF. For the other predictor - criteria relationships, differential predictability as described by a correlation was evaluated for each moderator variable subgroup. The strength of the associations were compared for significant differences (Edwards, 1984).

Selection procedures. Applicants for assembler positions came from regular enquiries from people seeking work and advertisements in the local community newspaper. On applying, individuals were asked to complete an application form and told they would be contacted for selection testing and interview in the near future.

When assemblers were required due to changes in production or the departure of

employees, individuals applying for jobs were asked to attend an interview and testing session of approximately two hours.

Apart from the application form, the main components of the selection procedure were the Flanagan Aptitude Classification Tests (Flanagan, 1959), a general knowledge test, and a short interview based on a standard set of questions (Appendix 4). The application form requested biographical data (name, birth date, marital status, etc.), employment history, and educational level. The Flanagan Aptitude Classification Tests were part of a 19 job element test battery designed to test job-related abilities such as Inspection, Mechanics, Coding, Memory, etc. The sub-tests used for assemblers were Inspection, Precision, and Coordination, with all test scores being converted to stanine scores according to test instructions (SRA, 1954). The general knowledge test asked 10 questions such as "who is the Prime Minister of New Zealand?", and "what voltage is the domestic power supply in New Zealand?" The test was scored out of 10 and was used to provide additional information about applicants (Appendix 4).

Once all the company selection procedures were completed, applicants were asked to complete the self-assessment form and the SDS. The self-assessment requests and the unmarked SDS were available to personnel staff although such information was not used in making selection decisions. All test scores, application form information, and interview data were assessed by the personnel manager. In her opinion, the best applicants for the particular positions available were offered jobs at a pay considered appropriate for the employee concerned. GRADE, the organisational employment decision, was based on a four level pay scale.

Results

The 61 assemblers were employed from an applicant pool of 149. Twenty five (17%) omitted a response to the request for a self-assessment of overall job performance (SAJOB, $n = 51$) amongst which were 10 employed assemblers. The homogeneity of responses to the self-assessment items described by Cronbach's alpha (.75) was considered adequate for the calculation of an average of self-ratings made on the various job aspects (SELFAV, $n = 61$). The correlation between SAJOB and SELFAV was .42 ($p < .001$). The median for each of the self-assessment measures was the same (4.0), the mean similar (SAJOB = 3.94, SELFAV = 3.93) and SAJOB had a slightly larger standard deviation ($SD = .645$) than SELFAV ($SD = .435$).

Fifty five new employees completed the induction course and received ratings. The induction course ratings were considered sufficiently homogeneous (Alpha = .77) to be averaged to form the induction course criterion INDUCT. Six employees were not required to do the course as they were considered to have had adequate previous work experience. Seventh week job performance ratings were available for 43 cases and were averaged (Alpha .91) to form the criterion PERFORM. The remaining 18 cases had left the organisation at that stage of their employment. The correlation between INDUCT and PERFORM was not significant ($r = -.01$, $n = 39$, NS).

Self-assessment and induction course performance. The impact of the organisation employment decision (GRADE) on INDUCT was not significant (Adjusted $R^2 = .05132$; $F = 3.380$, $df 1, 43$, NS). The effect of entering SAJOB was a nonsignificant R^2 change (.07805, $p > .06$), but a significant overall R^2 of .15094 ($F = 3.733$, $df 2, 42$, $p < .04$). When the average of self-assessment ratings (SELFAV) was used in place of SAJOB the R^2 change was not significant (.00394, NS), neither was

the overall R^2 with GRADE and SELFAV in the analysis ($R^2 = .06924$; $F = 1.934$, $df 2$, 52 , NS).

Self-assessment and appraised job performance. The impact of GRADE on PERFORM was not significant (Adjusted $R^2 = -.02756$; $F = 0.061$, $df 1$, 34 , NS). The R^2 change (.00153, NS) resulting from the entry of SAJOB was not significant, neither was the overall R^2 (.00332; $F = 0.055$, $df 2$, 33 , NS) at this final step. When SELFAV was used in place of SAJOB the analysis produced the same nonsignificant results.

Self-assessment and turnover. Two linear discriminant functions (LDF) were calculated, one using GRADE only, the second using GRADE and SAJOB. The analyses were compared in terms of classification efficiency and canonical correlations to establish the contribution made by self-assessment above that of the organisation selection decision. The first LDF using GRADE only was not significant with a canonical correlation of .29137 correctly classifying 36.7% of cases. The second analysis using GRADE and SAJOB computed two discriminant functions. The canonical correlation for Function One was .30043 and for Function Two .10915 with neither function reaching significance. The second analysis correctly classified 47.0% of the cases. Grade contributed 77.3% of the variance accounted for by Function One and 28.8% of the variance in Function Two. SAJOB was responsible for 22.9% of the variance in Function One and 73.5% of the variance in Function Two. Examination of the group means on the TURN criterion for SAJOB (functional turnover 4.0, SD .53; dysfunctional turnover 4.1, SD .66; still employed 3.8, SD .67) showed no major differences.

SELFAV was substituted for SAJOB and the analyses rerun. The canonical correlation using GRADE only was .32968 (Chi square = 6.559, $df 2$, $p < .04$). This first analysis

correctly classified 36.7% of the sample. The second analysis using GRADE and SELFAV computed two discriminant functions. The canonical correlation for Function One was .33955 and for Function Two .02175 with neither function reaching significance. The second analysis correctly classified 36.7% of the sample. Grade contributed 77.6% of the variance accounted for by Function One and 11.6% of the variance in Function Two.

SELFAV was responsible for 11.6% of the variance in Function One and 88.4% of the variance in Function Two. There were no major differences in the group means on the TURN criterion for SELFAV (functional turnover = 4.0, SD .28; dysfunctional turnover = 3.9, SD .48; still employed = 3.9, SD .45) although the variation in SELFAV ratings in the function turnover group was comparatively low. Whether self-assessment is represented by a overall job performance self-rating or an average of self-rated job items, it does not contribute significantly to the prediction of turnover in this study.

When the nine individual items making up the SELFAV measure were individually correlated with ratings on identical items made following the induction course and at the seventh week of employment, no significant associations were found.

Moderator analysis. No differences between moderator sub-group correlations were significant. Sub-group size was as low as three in some cases and for other moderators, such as job knowledge, missing data did not allow an evaluation of all relationships. The sub-group correlations and comparisons are shown in Appendix 5.

Discussion

Self-assessment did not contribute a significant level of additional predictive variance above that accounted for by organisational selection procedures on induction course ratings, job performance ratings, or turnover. This finding held when self-assessment was requested as a single item rating overall job performance or as an average of various work aspects. Examination of group means on the TURN criterion for both SAJOB and SELFAV showed little difference across groups for both variables. As demonstrated in other research using self-assessment in selection situations (e.g., Anderson et al., 1984; Pannone, 1984), inflation of self-assessments was apparent with the majority of ratings being centred on the four of a five point scale (5 high, 1 low).

The relationship between the overall self-assessment rating (SAJOB) and the averaged self-assessment ratings (SELFAV), $r = .42$ ($p < .001$), was not strong considering the variables were measuring the same construct. It may be that an applicant's perception of overall job performance is either different to or includes aspects of work other than those specified by the individual items making up SELFAV. It may also have been expected that specific aspects of the job such as timekeeping, tidiness, or use of tools may be more easily understood and more accurately rated than say an overall measure of a job that one has not yet done. In this investigation these individual measures were equally poor predictors of identical work aspects rated after induction and at the seventh week of employment. In fact, SAJOB as an overall self-assessment of work performance was the only self-assessment to come close to providing any predictive impact. The organisation's selection decision represented by GRADE did not have a significant impact on INDUCT. The entry of SAJOB into the analysis did not provide a significant R^2 change but did contribute enough variance to result in a significant overall R^2 in the prediction of INDUCT.

This study highlights the importance of how self-assessment requests are made. In this instance there were 10 applicants (16.4%) who completed self-assessments of individual aspects of work but omitted a response to the request for a self-assessment of overall work performance. This was despite pilot testing of the format with 15 individuals which revealed no problems. It may be that the arrangement of the request may have been responsible for applicants not completing it. The overall request may have been better presented in a more distinctive typeface surrounded by a greater proportion of blank space to ensure it was noticed and responded to (Sudman & Bradburn, 1983).

The large number of possible moderator sub-group comparisons, while increasing the chance of finding a significant result, allows a demonstration of consistency as far as moderator effects are concerned. Despite some seemingly large differences in sub-group associations, no significant moderator effects were detected (Appendix 5). In a small sample such as used in this setting the sub-group size available for analysis is often too small to have sufficient statistical power. In addition, some sub-groups often have no cases to analyse. Tantalisingly strong associations based on three cases having enactive job knowledge (e.g., SAJOB - INDUCT correlation, see Appendix 5) could mean a strong moderator effect exists, that the three cases involved for some other reason were highly accurate self-assessors, or that chance played a role in the result. Not having a larger group on which to test such possibilities prohibits any definite conclusions. Statistically in the present case, it appears that all the sub-groupings of these New Zealand assemblers are equally inaccurate assessors of work performance criteria.

It must be noted that the performance criteria utilised in applied settings are not necessarily accurate measures of work performance just because an organisation uses them. In the present investigation the relationship between two organisational measures of work performance, INDUCT and PERFORM, was not significant. Considering the induction course was based directly on the skills required on the job and that INDUCT and PERFORM ratings used the same format and content, it is surprising that the measures were not more closely related. In addition, the selection decision GRADE showed no evidence of predictive validity. Whether the failure of self-assessment to predict performance criteria is due to poor predictive characteristics or a lack of valid performance measures is difficult to establish. In this particular setting, the observed relationships between organisational measures allows little confidence in the criteria available.

CHAPTER SEVEN

EVALUATION OF SELF-ASSESSMENT VALIDITY IN A FOOD PROCESSING COMPANY

This setting was part of an organisation which processes, cans, and freezes seasonal fruit and vegetables. The organisation employed around 700 permanent employees and an additional 500 seasonal employees. The two year period over which the investigation was conducted involved two consecutive fruit and vegetable processing seasons. Each season was treated as a separate study because the job content of the seasonal worker positions changed substantially from the first to the second season. Study A was conducted during the first season and Study B in the second.

Apart from examining the addition to predictive variance offered by self-assessment to on-the-job performance, absence, and turnover criteria, Study A investigated the moderating effects of gender, age, previous job experience, education, job knowledge, and self-assessment experience. In addition, two hypotheses were considered in Study A. The first hypothesis considers the effects of different types of job information on the accuracy of self-assessment as suggested by Bandura's (1977b) self-efficacy theory (see Chapter Three). The second hypothesis is an evaluation of Mabe and West's (1982) contention that self-assessment requests couched in comparative terms will be more accurate than requests not asking for comparative self-assessments (see Chapter Four). The hypotheses can be stated thus: -

1. Applicants who are able to base their self-assessments on enactive job information will be more accurate self-assessors than applicants who have had access to vicarious job information, who in turn will be more accurate than

those having access to judgmental information. Applicants having no access to job information will be the least accurate self-assessors.

2. Self-assessment information requested in relative or comparative terms should be more accurate than information requested in absolute, non-relative terms.

Study A

Method

Subjects

The subjects were applicants who accepted an offer of a job as a seasonal worker for a food processing company during the first three months of the food processing season ($n = 176$). The average age was 25.9 months, ranging from 14 to 60 years; 135 (76.7%) were female and 41 (23.3%) were male. As applications for work were made in November and the majority of seasonal positions were not available until January of the following year, only a small proportion (approximately 20%) of total applicants were employed. The main reason applicants did not take jobs was because they had other commitments by the time offers of work were made. According to the employing organisation, most available applicants were offered positions at some point during the season.

Procedure

Development of self-assessment requests. Seasonal workers were employed in a variety of unskilled positions to assist permanent staff with the processing of harvested fruit and vegetables. To establish which jobs could be regarded as seasonal, company personnel were interviewed, line workers observed, and organisational documentation and union agreements were reviewed. From this information seven

types of work were identified and brief descriptions compiled and checked with workers, supervisory, and personnel staff. The descriptions used are listed below.

Inspection. Inspects fruit, vegetables, or cans as they pass along a belt. Fruit or vegetables that are marked or damaged are removed along with any foreign matter such as sticks or leaves. Inspection of cans involves removing damaged or poorly labeled cans.

Trimming. Fruit or vegetables are taken from a moving belt and trimmed to remove marked or damaged parts.

Belt work. This involves working on the production line and doing a number of things by hand. Some examples of this are painting carton flaps with glue; filling cans, cartons, or packets with fruit or vegetables; labelling cans; and putting cans into crates.

Machine feeding. Feeding fruit or vegetables into processing machines. For example, placing apples or pears on the feed cups of peeling machines.

Cleaning. Carrying out general cleaning, tidying up, and washing down duties.

Machine operating. Operating machines which process fruit or vegetables. Some machines involved are those which close cartons, seam cans, bag frozen vegetables, put cans on pallets, and control the flow of juice fruit and vegetables.

Heavy line work. This involves such work as carrying, stacking, unloading bags, and cartons, sewing large bags, emptying bins or cases, putting together cardboard cases, etc.

Development of the self-assessment request was based on the same guidelines described and discussed in the investigation conducted in the electronics manufacturing setting (see Chapter Six). These guidelines are briefly summarised below.

- a) No job information in addition to that normally available to applicants during the selection procedures was provided by the self-assessment request.
- b) Self-assessment requests were self-administered.
- c) Applicants were asked for an overall self-assessment of future work performance.
- d) Self-assessment requests were made as the last stage in the selection process prior to the applicant being informed of any offer of employment.

The descriptions of the different types of work were incorporated into a self-assessment request form and administered to 15 employed workers to check readability, understanding, and acceptability. As five of the workers had slight difficulty with understanding some of the words in the self-assessment request, changes to simplify the wording were made.

The self-assessment request asked applicants to rate how they thought they could do the seasonal work. The scale on which self-ratings were to be made was a 5 to 1 scale used by the company for rating job performance (5 = an excellent standard of

performance, 1 = a poor standard of performance). The self-assessment requests were incorporated into the company seasonal worker application form. At the time of application it was not known which of a number of seasonal jobs applicants would be employed to do. Self-assessments were therefore requested for all seven seasonal jobs. A copy of the application form can be seen in Appendix 6.

Criteria. Criteria collected in this study were appraised job performance (PERFORM), turnover (TURN), and absence (AB-DURATION). To assist in future employment decisions each seasonal worker was given an overall rating of job performance by their supervisor when they left. The reliability of this rating was measured by comparing overall performance ratings given by supervisory staff and ratings given by personnel staff on a sample of 65 seasonal workers. The correlation between the two sets of ratings was .89 ($n = 65$). All ratings were made on the same scale used for ratings of work suitability (5 = excellent, 1 = not acceptable).

In the present study workers were assigned into either of two turnover categories. "Stayers" were defined as those workers who stayed with the company until their services were no longer required. These individuals then either had their seasonal contracts terminated or were transferred to permanent staff in other sections of the company's operations. "Dysfunctional turnovers" were those workers who stopped work before their seasonal contract was completed. The reasons for leaving in this second group included, abandoning job, resigning due to family responsibilities, leaving the district, and ill health. There were no workers in this sample who could be classified as "functional turnovers" (e.g., dismissed, asked to resign, etc.).

The measure of absenteeism used was the number of days on which a worker was absent and had not contacted the company to give a reason for the absence (AB-

DURATION). Such absences were regarded by personnel staff and line supervisors as representing an unacceptable attitude to work. They did not know when, if ever, a worker would return and so could not adjust work force levels. As absenteeism figures were used in pay calculations and had to be agreed on by the company and the worker, this absence data was felt to be accurate. AB-DURATION was coded as the number of days workers were absent with no reason.

Moderator variables. To investigate possible moderator effects subject information was collected on gender, age, job knowledge, and self-assessment experience as in the electronic manufacturing investigation (see Chapter Six). In addition, education was investigated as a moderator. Education was measured as the secondary school form level reached.

To obtain an indication of the effect of comparative versus non-comparative requests for self-assessment information, two versions of the self-assessment form were used. The non-comparative version (NONCOM) asked for self-assessments with no reference to a comparison group. No social comparison terminology such as "average" or "better than most" was used in the wording. The comparative version (COMP; see Appendix 6) was identical, with the exception of two sentences which put the request for self-assessment in comparative terms (e.g., "when answering those questions which ask for a rating, compare yourself with the sort of people you think would be doing this type of work"). The order in which the two versions were distributed to applicants was randomised.

Analysis. Regression analyses were used to examine the impact of self-assessment on performance, turnover, and absence criteria, in addition to that accounted for by the organisational selection process. Potential moderators of the

relationship between self-assessment and the different criteria were considered using moderated regression analysis (Arnold, 1982).

Selection procedures. The company opened a register of seasonal workers at the beginning of November, being the early stages of New Zealand's fruit and vegetable harvesting season. Applicants completed an application form which included self-assessment requests after speaking briefly to company personnel staff to establish availability and telephone contact. The application form requested biographical (name, birth date, address, etc.), employment history information, and self-assessments of job performance. Personnel staff then rated the applicant on a five point suitability for employment scale (1 = excellent, 5 = unacceptable) based on what they knew from the application form, impressions of the person, and any records of previous experience with the company. Applicant record cards were filled out for each candidate and filed in a "potential seasonal workers" file. As production requirements dictated, the personnel office would be informed by line supervisors when seasonal workers were required. Starting with those applicants rated as "1", personnel staff would work through the potential seasonal worker file contacting applicants until the worker requirement had been filled. As there were severe shortages of seasonal workers during the season, a number of attempts were made to contact all applicants. This meant that even individuals rated as unacceptable, in terms of work suitability, were employed by the company.

Results

Correlations were calculated between the three types of criteria, rated job performance (PERFORM), turnover (TURN), and absence (AB-DURATION) with only the relationship between PERFORM and AB-DURATION being significant ($r = -.3015$, $p < .001$). The correlation between the two main predictors of interest, self-assessment (SAJOB) and suitability for employment rating (SFE) was also significant ($r = .1619$, $p < .02$). The moderating effect of job knowledge and comparative (NONCOM/COMP) self-assessment requests was tested together with other potential moderators (gender, age, job knowledge, education, and self-assessment experience) and is reported below for each criterion.

Self-assessment and appraised performance: The impact of SFE by itself, on PERFORM, was significant (Adjusted $R^2 = .02896$; $F = 6.219$, $df 1, 174$, $p < .02$). The change in R^2 (.02268) obtained by entering SAJOB was also significant ($p < .05$) giving an overall adjusted R^2 of .04629 ($F = 5.24674$, $df 2, 173$, $p < .01$). SAJOB was related to PERFORM in a negative direction (BETA = $-.15260$, $p < .05$).

Following the entry of SFE and then SAJOB, the seven variables being considered as moderators (gender, age, previous job experience, education, NONCOM/COMP, job knowledge, and self-assessment experience), were entered into the analysis to partial out the main effects. The last step in the analysis was the entry of the seven variables which were the products of the potential moderators and self-assessment experience. With all variables entered, the overall multiple R was .50172 (Adjusted $R^2 = .17643$; $df 16, 159$, $p < .0001$). Figure 7.1 shows the significant relationships obtained at the final step of the analysis.

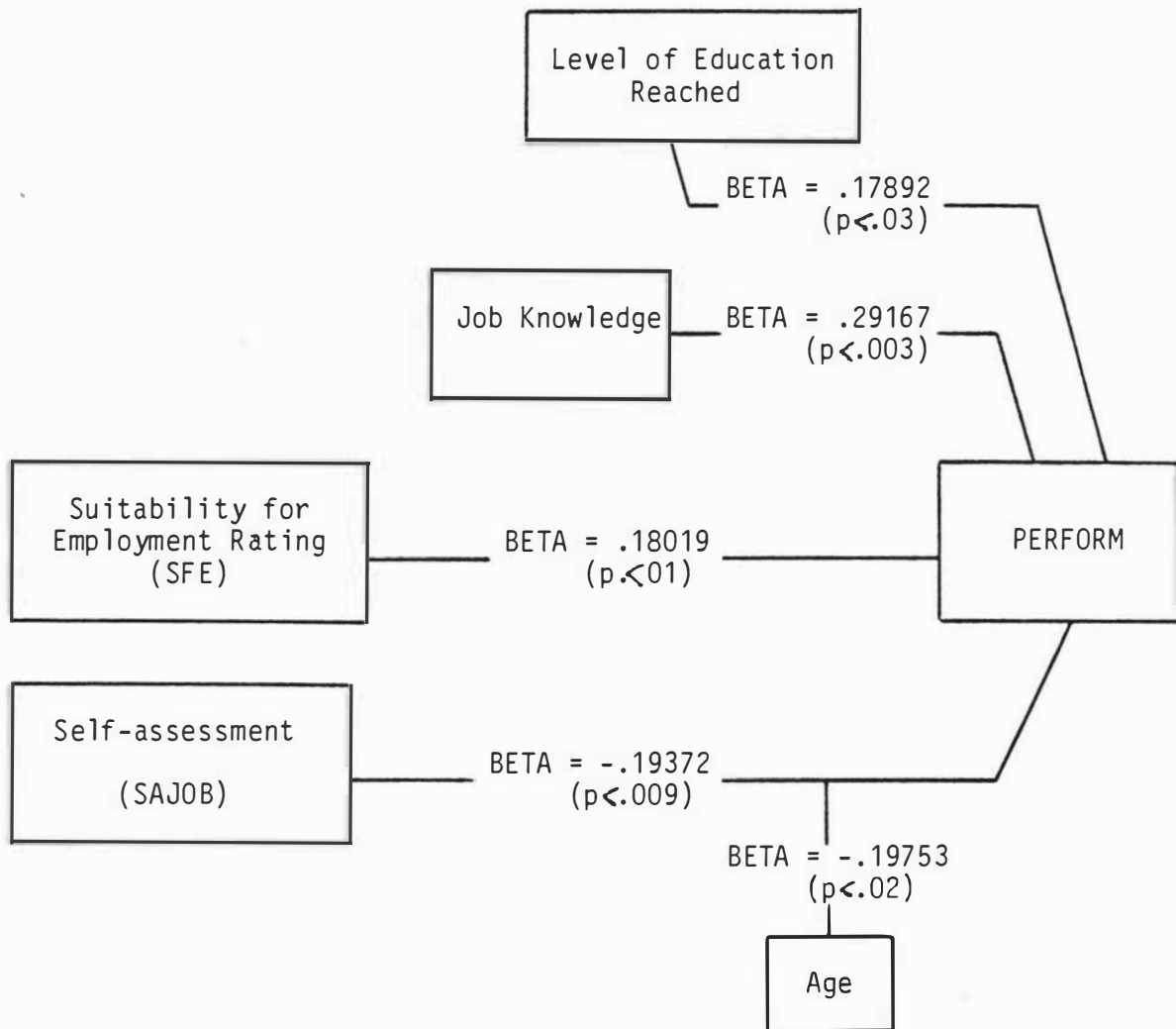


Figure 7.1. Moderated Regression Results for Appraised Performance Criteria.

Of the seven potential moderators only age showed a significant effect (BETA = $-.19753$, $p < .02$). The negative BETA shows that as Age increases, the slope of the SAJOB - PERFORM relationship decreases. To illustrate this moderating effect, the sample was divided into quartiles based on age (Table 7.1).

The correlation between SAJOB and PERFORM was calculated for each group. Although no correlation was significant, it can be seen from Table 7.1 that the

direction of the relationship was similar for three of the younger groups but was in an opposite direction for the oldest group. In other words, as age increased the negative slope decreased to the point where the older age group demonstrated a positive relationship.

Table 7.1

Correlations Between Self-Assessment and PERFORM for Four Age Quartile Groups

	Age Range	Number	Percent	r	Prob
1	14 - 18 yrs	n = 46	26.1%	-.283	p>.056
2	19 - 21 yrs	n = 42	23.9%	-.215	p>.171
3	22 - 30 yrs	n = 43	24.4%	-.220	p>.156
4	31 - 60 yrs	n = 45	25.6%	.272	p>.070

At this last step of the analysis a number of main effects were also significant. SFE and SAJOB were both significantly related to PERFORM, SFE in a positive direction, SAJOB in a negative direction. Examination of a cross-tabulation of SAJOB and PERFORM scores suggested that the negative relationship was due to a lack of cases giving themselves low assessments and performing well and a large number of cases rating themselves as good performers and receiving low performance ratings. SAJOB ratings were skewed to the positive end of the scale whereas PERFORM ratings were more normally distributed.

Job knowledge was positively related to PERFORM (Figure 7.1). The mean of PERFORM (5 = excellent, 1 = unacceptable) for each of the categories of job knowledge were, enactive 3.35 (SD = 0.94), watching others 3.31 (SD = 1.01), being

old 3.43 (SD = 0.88), and no source 2.57 (SD = 1.06). There was little difference, in terms of work performance, between cases with enactive experience and cases with observational or other job knowledge. Those cases with no source of job knowledge were, on average, rated lower.

The impact of education on PERFORM (Figure 7.1) suggested that the higher the level of secondary education achieved, the more likely workers were to receive a high performance rating.

Self-assessment and turnover. The sample was divided into two groups. The 82 (46.6%) cases leaving at the end of the processing season were coded two and labelled "stayers". The 94 (53.4%) cases leaving before the end of the season were coded one and called "dysfunctional turnovers". The impact of SFE on TURN was not significant. The change in R^2 obtained by entering SAJOB was significant (R^2 change = .04828, $p < .007$), giving an overall adjusted R^2 of .04614 ($F = 5.23228$, $df 2, 173$, $p < .006$). Following the entry of all variables (gender, age, previous job experience, education, NONCOM/COMP, job knowledge, and self-assessment experience), the overall multiple R was .48386 (adjusted $R^2 = .15705$; $F = 3.03771$, $df 16, 159$, $p < .0002$). The pattern of relationships is shown in Figure 7.2.

The negative relationship between SAJOB and TURN shows cases rating themselves highly were more likely to engage in dysfunctional turnover and those rating themselves lower were more likely to stay to the end of the season. A cross-tabulation of cases between SAJOB and TURN indicates that this positive relationship is partly due to self-assessment ratings being skewed to the high end of the scale while turnover categories were reasonably evenly distributed. On the five point scale

(5 = excellent, 1 = poor), the average self-rating of the stayers was 3.82 (SD = .78), for dysfunctional turnovers 4.11 (SD = .91).

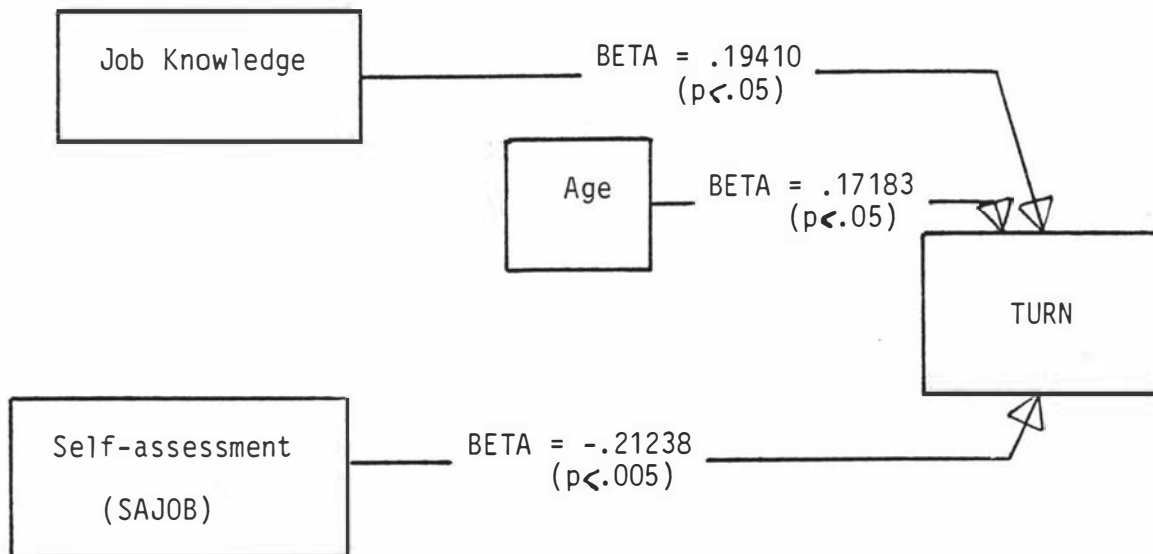


Figure 7.2. Moderated Regression Results for Turnover Criteria.

No moderator effects were detected for TURN although job knowledge and age had a significant main effect (Figure 7.2). Workers with job knowledge were more likely to leave before the end of the season (Figure 7.2). The major differences between dysfunctional and stayer groups can be seen only when enactive and no source job knowledge categories are compared. Of cases in the dysfunctional turnover category, 36 (38.3%) had no source of job knowledge and an equal number had enactive experience. In comparison, 13 stayers (15.9%) had no source of job knowledge, whereas 47 (50.0%) had enactive experience. Cases having job knowledge other than that involving enactive experience were equally likely to be classified as either dysfunctional or as stayers.

The major reasons for leaving amongst the 94 workers engaging in dysfunctional turnover were resigning due to domestic responsibilities ($n = 18, 19.1\%$), abandoning work ($n = 13, 13.8\%$), resigning for no given reason ($n = 23, 24.5\%$), resigning to go to a better job ($n = 10, 10.6\%$), and resigning to return to educational studies ($n = 13, 13.8\%$).

The negative main effect for age shown in Figure 7.2 indicates that older workers were more likely to stay until the end of the season, or alternatively, the younger workers were more likely to leave before the end of season. The average age of stayers were 28.8 years compared to dysfunctional turnovers which was 23.4 years. Splitting the sample at the median age (21.5 yrs) showed that the younger group were more likely to leave for university or other study whereas older workers were more likely to leave for domestic reasons.

Self-assessment and absence. AB-DURATION was coded as the number of days workers were absent without a telephoned reason. The impact of SFE on its own, on AB-DURATION was not significant (Adjusted $R^2 = -.00206$; $F = .64089$, $df 1, 174$, NS). The change in R^2 (.00665) obtained by entering SAJOB was not significant, nor was the overall multiple R (.34976) at the final stage of the moderator analysis (adjusted $R^2 = .03401$; $F = 1.38508$, $df 16, 159$, NS). There was, however, a significant moderating effect demonstrated by age (BETA = $-.327709$, $p < .0003$). To illustrate this effect the sample was divided into quartiles and correlations calculated for each as shown in Table 7.2.

For the first quartile age group (14 - 18 yrs), the higher the level of SAJOB the higher the level of unexplained absence, whereas for the older workers in the fourth

quartile age group (31 - 60 yrs) the direction of the relationship was reversed (Table 7.2).

Table 7.2

Correlations Between Self-Assessment and Absence for Four Age Quartile Groups

	Age Range	Number	Percent	r	Prob
1	14 - 18 yrs	n = 46	26.1%	.315	p<.039
2	19 - 21 yrs	n = 42	23.9%	.083	p>.600
3	22 - 30 yrs	n = 43	24.4%	.183	p>.240
4	31 - 60 yrs	n = 45	25.6%	-.339	p<.023

Discussion

Self-assessment (SAJOB), in the presence of the organisation's job suitability rating was significantly related to appraised job performance (PERFORM) and turnover (TURN), but not to absence (AB-DURATION). Job suitability ratings (SFE) demonstrated a significant impact on PERFORM only. Neither Hypothesis 1 in respect of job information, nor Hypothesis 2, which dealt with non-comparative versus comparative requests (NONCOM/COMP), was supported. Moderator effects due to age were detected on PERFORM and AB-DURATION criteria. Some main effects were observed.

Hypothesis 1, which stated that the level of job information an applicant possessed would moderate the relationship between self-assessment and work performance criteria, was not supported. Seasonal workers who had reported enactive job

experience gave SAJOB predictions that were no more accurate than predictions given by workers with no reported job experience. Although judgments of one's potential performance are functionally related to subsequent performance, Bandura (1982b) notes that a number of factors can affect the strength of this relationship. These include "faulty self-knowledge, misjudgment of task requirements, unforeseen situational constraints on action, disincentives to act on one's self-percepts of efficacy, ill defined global measures of perceived self-efficacy, or inadequate assessments of performance, and new experiences that prompt reappraisals of self-efficacy" (p. 129).

The demand characteristics of the personnel selection setting is said to bias responses toward social desirability (Ash, 1980). It has been suggested that the tendency to be overly positive should be expected primarily from those who are least qualified in a situation where they have a vested interest (Wicklund & Gollwitzer, 1983). For example, when the purpose of a self-appraisal has been varied between research and administration (e.g., grading, selection), administrative self-assessments have been far more lenient than those being used for research (Farh & Werbel, 1986). Research on personality questionnaires has shown that job applicants return higher scores on distortion scales than similar people who are not undergoing selection (Elliot, 1981). Job applicants have also been shown to inflate self-assessments of experience and training on job related tasks (Anderson et al., 1984). It may have been that the disincentives present in this particular selection situation swamped any signs of a job information effect.

Hypothesis 2 stated that self-assessment information requested in relative terms would be a more accurate predictor than the same information requested in non-comparative terms. Although it is suggested that self-evaluation measures which require relative comparisons are better approximations of the actual meaning of criterion measures of

ability than absolute measures (Mabe & West, 1982), both approaches to collecting self-assessment information were similarly inaccurate predictors of subsequent job performance. People first attempt to evaluate their abilities on the basis of physical, nonsocial reality and if such means are unavailable, they make an evaluation through a process of comparison with other individuals (Festinger, 1954). Because little in the way of physical or nonsocial information was available to applicants (e.g., standards and descriptions of performance), including those who were past employees, self-assessments were more likely to be based on how an applicant compared her or himself to others. It may be that in this case self-assessments, whether couched in comparative or absolute terms, were treated in a similar way.

The relationship of SAJOB to PERFORM was in the opposite direction to that which was expected. Applicants rating themselves as good performers were more likely to be rated as poor performers on the job compared to those applicants who did not rate themselves as highly. In this investigation applicants displayed a marked reluctance to use scale points describing their future performance as less than "reasonable". The distribution of self-ratings skewed to the "excellent" end of the scale suggested that the particular situation was a disincentive for applicants to make anything other than positive self-ratings of overall work performance. Applicants were behaving in a manner that best suited their purpose, possibly reflecting the demand characteristics of the selection situation.

The negative relationship demonstrated between self-assessment and turnover indicated that those who rated themselves as good performers were more likely to leave the job before the season was complete. Ekpo-Ufot (1976) also found that ratings of self-perceived abilities relevant to a task (SPART) were related to turnover. Ekpo-Ufot's finding was "interpreted to mean that the higher the worker perceives his standing on

task relevant abilities the greater the probability of his remaining on the job" (p. 412), a reversal of the relationship shown in the present study.

Serious turnover problems have been a reported feature of seasonal worker groups. Dunnette and Maetzold (1955) state, "it is not uncommon for an applicant to state on his application blank that he will work all summer only to leave with no apparent reason after working only a short time" (p. 308). Why those in the dysfunctional turnover category generally had higher self-ratings than those who were classified as stayers is difficult to assess. One clue may lie in Korman's (1970) hypothesis of work behaviour that suggests individuals will engage in and find satisfying those behavioural roles which will maximise their sense of cognitive balance or consistency. It is possible that those workers who rated themselves highly were more likely to find that the unskilled nature and the repetitive character of the work was intolerable. This result may also be due to the inflation affected distribution of self-ratings when compared to the relatively even distribution of cases between stayer (46.6%) and dysfunctional turnover (53.4%) categories.

There is not a lot of applied research considering age as a moderator in self-assessment/criterion relationships. Brief, Aldag, and van Sell (1977) report a nonsignificant difference between age subgroup (< 29.9 yrs and > 29.9 yrs) correlations between self-and superior evaluations of job performance. Other research has reported a similar lack of moderator effect due to age (DeNisi & Shaw, 1977). Lawler (1967) found in a managerial sample that older managers (over 40 yrs) were more lenient than younger managers (under 40 yrs), when their self-ratings were compared to supervisory performance ratings.

The moderating effect of age on the SAJOB - PERFORM relationship (Figure 7.1) found in this seasonal worker sample suggests that older applicants (31 - 60 yrs), when compared to the other age groups, may have responded differently to the request for self-assessment ratings. Although the moderating effect was not significant in terms of subgroup correlations (Table 7.1), the change in the direction of the relationship across subgroups is certainly worthy of further investigation.

Age was also identified as a moderator on the absence criterion. The negative relationship (BETA = $-.32771$, $p < .0003$) indicates that as age increases the slope of the regression line of the SAJOB - AB-DURATION relationship decreases. Splitting the sample into age group quartiles showed for the youngest age group SAJOB was significantly related to AB-DURATION in a positive direction, but for the older age group (31 - 60 yrs) the significant relationship was negative (Table 7.2). It appears that for the younger age group, the higher the SAJOB the greater the amount of AB-DURATION. For the older age group, the higher the SAJOB rating the lower the frequency of AB-DURATION. Older seasonal workers who rated themselves as good workers were more likely to attend work, or give satisfactory reasons for not attending, than those who rated themselves as less competent workers. No obvious explanation can be offered for this trend apart from the possibility that it was due simply to a tendency for older workers not to inflate self-ratings. Both AB-DURATION and TURN measures showed that older workers who rated themselves higher tended to perform in a manner that was beneficial to the organisation. Similar ratings made by younger workers meant the opposite.

Job knowledge demonstrated a main effect on PERFORM and TURN. Workers with more experiential job knowledge being more likely to be rated as good workers whereas those with less job knowledge were more likely to get poor job performance

ratings. The relationship between job knowledge and TURN (Figure 7.2) indicates that those staying for the season were more likely to have enactive job knowledge.

Although the work done by seasonal workers does not require a high level of skill, it may be that having enactive experience helps seasonal workers to "survive" the season.

The impact shown by the level of education on PERFORM (Figure 7.1) suggests workers with more education were more likely to receive higher PERFORM ratings while workers with less education were more likely to receive lower PERFORM ratings. Personnel staff reported that university students were generally very good workers as "they need the money". While there were only 14 (7.4%) students in the sample, too few for separate analysis, the performance of this particular sub-group may have been partly responsible for this relationship.

Age was negatively related to TURN. In this particular sample it seems that the older workers were more likely to stay as long as the company required, whereas younger workers were apt to leave before the season's end. The evidence for permanent work, although not consistent, suggests the reverse. Rhodes (1983), for example, showed the older workers were more likely than their younger counterparts to turnover. The reason for the conflicting results shown in this study is not obvious. It may be that over a brief period of employment, as is common in seasonal work, older workers are more likely to stay for the entire contract period. Over longer periods of employment it is the younger workers who have the longer periods of tenure.

In conclusion, it appears that the selection situation had a major effect on the relationship between SAJOB and PERFORM, a relationship moderated by age in the

case of two criteria. Older workers (>30 yrs) appeared less likely to inflate overall self-assessments of performance resulting in positive relationships with performance appraisals and lower rates of unexplained absences.

The generally weak relationships between self-assessment and the various work performance criteria may have been affected by the way applicants used the self-assessments. The seasonal work in this organisation is largely machine paced, unambiguous, routine, methodologically invariant, and closely supervised. There are, however, differences between the major tasks. Some are more popular than others. It is possible that applicants, when asked for ratings of proficiency, are really expressing preferences for work types rather than giving ratings of ability. High self-ratings on some jobs may be seen by applicants as enhancing their chances of being offered work in a desired area. Conversely, low self-ratings may be seen as assisting applicants to avoid work on less desired tasks. Such a trend would be demonstrated when a preference for a particular job is positively related to a self-assessment of performance on that job.

Personnel staff dealing with seasonal applicants reported that although all applicants were able to complete the application form which included the self-assessment request, a number required assistance with reading the instructions and understanding the job descriptions. In addition, personnel staff felt applicants were initially "put off" by the large amount of written material in the application form (3 pages). As a result subsequent self-assessment requests were designed to use simpler language and ask for less information.

Study B

Study B was conducted during the following growing season. As well as evaluating the validity of self-assessment, this second investigation considered the question of job preference having an effect on self-ratings as raised in Study A. Specifically this involved describing the relationship between an applicant's preference for a job and his or her self-assessment and establishing whether or not job preference offered any predictive variance over and above that accounted for by self-assessment. The role of age as a moderator of self-assessment - job criteria relationships as demonstrated in Study A (George & Smith, in press) was further investigated in addition to the moderating effect of gender, secondary school qualifications, and source of job knowledge.

Method

Subjects

Compared to Study A, the number of employed seasonal worker applicants ($n = 375$) increased due to the expanding role of seasonal workers in this particular organisation and extending the data collection period from three months to the entire season (eight months approximately). The average age was 26.8 years, ranging from 15 to 60 years old, 285 (76%) were female, 90 (24%) were male. Sixty subjects had secondary school qualifications obtained in the sixth or seventh Forms, 78 had School Certificate, and 237 reported no secondary school qualifications.

Procedure

Development of self-assessment requests. Functional changes to improve production efficiency in the organisation resulted in changes to the general categories of seasonal work from those used in Study A. Two new categories of work were

introduced for seasonal workers. Firstly, approximately 40 applicants were to be employed as field workers to do the manual work required during the planting and harvesting of the tomato crop. Secondly, due to an increased level of fresh fruit and vegetables for processing being obtained from outside the immediate region, some seasonal drivers were also required. The exact number of seasonal workers to be used as drivers depended on the availability of qualified applicants and local versus other region crop yields. The number required was not expected to exceed 20, but this category was included as it involved an identifiably different class of work.

Other changes in the organisation of seasonal tasks required the dropping of self-assessment requests for some categories and a change in the nature of others. Belt workers previously involved only in packing and labelling were to be involved in a range of trimming, inspection, and feeding on activities. These categories were therefore combined under the title belt work. The heavy line work category became the responsibility of permanent staff so this category was dropped for seasonal workers. Cleaning, previously a separate job category, had become part of a general line work category which also involved the carrying, stacking, and filling of bags, cartons, and bins. The tasks in this category required no heavy lifting. The machine operating category remained unchanged from Study A. As a result of these changes self-assessment requests were made for five main categories of work; driving, belt work, machine operating, field work, and general line work.

The format of the application form requesting self-assessment also was changed. Personnel staff reported that during Study A some applicants required assistance with reading the instructions and understanding the job descriptions. To avoid similar problems in Study B, registration procedures were changed so some biographical data was requested verbally and recorded directly onto record cards. The remaining

information required by the application form was reduced. Written job descriptions were simplified and accompanied by a 30 mm by 60 mm sketch to reduce the amount of text and to further describe the work involved (Appendix 7).

To measure preferences applicants may have had for the five different types of seasonal work, they were asked "If you had a choice of jobs among driving, belt work, machine operating, field work or general line work, which job would be; your 1st choice.....; your 2nd choice.....; your 3rd choice.....; or your 4th choice.....?" (Appendix 7). The work type not mentioned in any of the four choice categories was treated as a fifth choice. The first choice was coded five, the second choice as four, through to the fifth choice which was coded as one.

Criteria. Job performance and turnover criteria were the same as those used in Study A. Supervisors appraised the job performance of each seasonal worker on termination using the same five point scale used for rating job suitability (5 = excellent, 1 = not acceptable). For the turnover criterion the sample was classified into "stayers" as workers who stayed until no longer required (n = 187) and as "dysfunctional turnovers" as workers who left before the completion of their season contract (n = 186). There were two workers who were dismissed during the season. These were the only functional turnovers (Dalton, Krackhardt, & Porter, 1981) and were deleted from the sample for the analysis involving the turnover criteria.

The absenteeism criterion used in Study A was the number of days workers were absent from the job for no particular reason. As discussed in Chapter Two, a common alternative to such duration measures of absence is a frequency index which is based on the number of absences in a given period rather than the amount of days or shifts missed (Chadwick-Jones, Nicholson, & Brown, 1982). It was decided to use

both a duration and frequency measure in this study to see if there was any difference in the predictive validity of self-assessment for each measure. AB-DURATION was defined as the number of days an employee was absent from work with no telephoned reason. AB-FREQUENCY was defined as the number of times (irrespective of length) an employee was absent from work with no telephoned reason.

Moderator variables. To investigate possible moderator effects, subject information was collected on age, gender, educational qualifications, and source of job knowledge. With the exception of educational qualifications, moderators were the same as those used in Study A. Instead of requesting form level reached as an indication of education level, personnel staff indicated that they would prefer to have applicants state the type of secondary qualification attained. Often no documentary proof (such as a certificate) was available for form level reached and the staff felt a number of applicants misrepresented this information. Secondary school qualifications were standard across all secondary schools and could be supported by a certificate. Education was therefore defined as the level of secondary qualification obtained.

Selection procedures: The selection procedures in Study B were the same as for Study A with the exception of the changes mentioned above. Personnel staff rated the applicants on a five point suitability for employment scale (SFE) on registration. Workers were contacted, offered employment, and placed as production requirements dictated. As in Study A, there was a shortage of workers during the season and a number of attempts were made to contact all applicants. This was reflected in the employment of 18 applicants rated as below average and three rated as unacceptable.

Results

The relationships between the four criteria, appraised job performance (PERFORM), turnover (TURN), duration of unexplained absence (AB-DURATION), and frequency of unexplained absence (AB-FREQUENCY), are shown in Table 7.3.

Table 7.3

Correlations Between Appraised Performance, Turnover, and Absence Criteria (duration and frequency)

	PERFORM	TURN	AB-DURATION	AB-FREQUENCY
PERFORM	.00.	.37 **	-.30 **	-.22 **
TURN	.37 **	.00	-.06	-.01
AB-DURATION	-.30 **	-.06	.00	.68 **
AB-FREQUENCY	-.22 **	-.01	.68	.00

** = $p < .001$ (2 tailed)

A similar negative relationship as seen in Study A between PERFORM and AB-DURATION was found in Study B ($r = -.30$). The relationship between PERFORM and AB-FREQUENCY in Study B was also negative but not as strong ($r = -.22$). The positive relationship between PERFORM and TURN suggests that stayers tended to be rated as higher performers. The correlation between the suitability for employment rating (SFE) and self-assessment (SAJOB) was negative ($r = -.14, p < .01$). Eighty four percent of SAJOB ratings were above the mid point of the scale. In comparison to Study A, applicants appeared more likely to use the full range of the scale with seven applicants rating themselves as poor.

Applicants were asked to make ratings on the five types of work available to seasonal workers. The average self-rating on these job types were, driving 3.2 (n = 352, SD 2.6), belt work 7.3 (n = 372, SD 2.2), machine operating 6.5 (n = 366, SD 2.6), field work 4.8 (n = 352, SD 2.9), and general line work 6.7 (n = 365, SD 2.7). A varying number of applicants did not complete self-assessment for some categories. Personnel staff reported that in most of these cases applicants did not think they could do those particular jobs. Of the 375 in the sample nine (2.4%) were hired as drivers, 246 (65.6%) as belt workers, 13 (3.5%) as machine operators, 16 (4.3%) as field workers, and 91 (24.3%) as general line workers. The two major areas of employment, belt work and general line work accounted for 89.9% of seasonal staff and had the highest average self-assessment.

The impact of SAJOB on each of the job performance criteria was tested as part of a moderated regression analysis. As the first step, SFE was entered followed by SAJOB to test the additional predictive utility offered by self-assessment above that of the current organisational selection procedures. To test potential moderator effects, age, gender, educational qualifications, and source of job knowledge variables were entered next. The final step involved entry of the products of the moderator variables.

Self-assessment and appraised performance. The impact of SFE only on PERFORM was significant (Adjusted $R^2 = .08640$; $F = 36.369$, $df 1, 373$, $p < .0001$). The change in R^2 following the entry of SAJOB was not significant. Figure 7.3 shows the significant relationships at the last stage of the analysis.

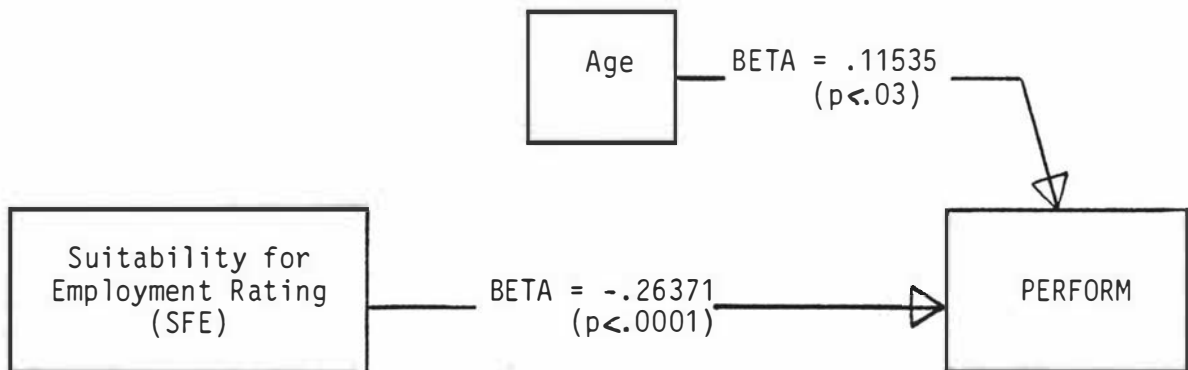


Figure 7.3. Moderated Regression Results for PERFORM (Appraised Job Performance).

At this final step the multiple R was .34659 (Adjusted $R^2 = .09595$; $F = 4.9695$, df 10, 364, $p < .0001$). No moderator effects were significant. Interestingly the impact of SFE was negative suggesting those applicants rated high during selection were more likely to be rated low performers on the job. Age demonstrated a significant effect on PERFORM indicating older workers were more likely to be rated high although this effect was not strong.

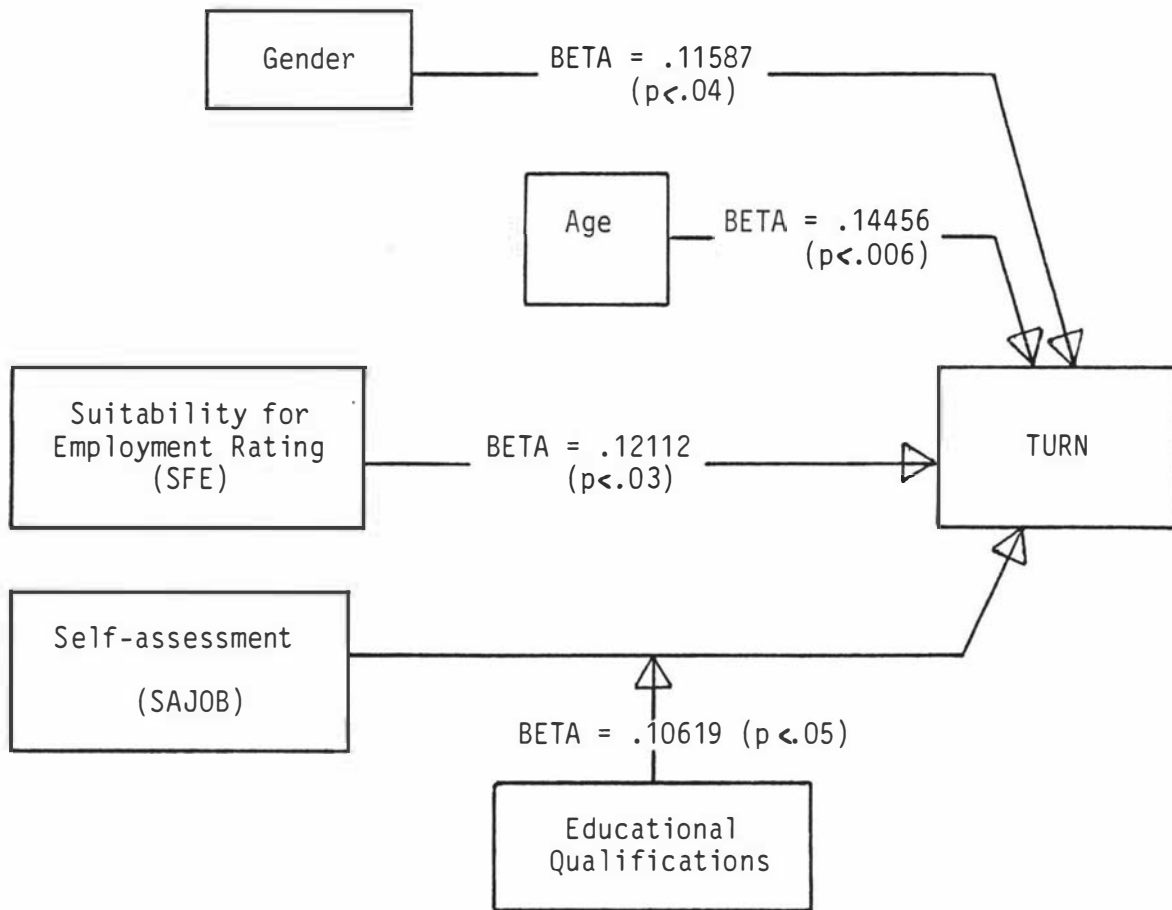


Figure 7.4. Moderated Regression Results for TURN Criteria.

Self-assessment and turnover. The sample was split into two similar sized groups being stayers ($n = 187$, coded 2) and dysfunctional turnovers ($n = 186$, coded 1). The impact of SFE on TURN was not significant, although in the presence of the other variables, SFE demonstrated a significant impact. The change in R^2 obtained by entry of SAJOB was not significant. Following entry of all variables the multiple R was .28214 (Adjusted $R^2 = .05418$; $F = 3.1309$, $df 10, 362$, $p < .0007$). Figure 7.3 shows that at the final stage of the analysis Age had a significant moderating effect on the impact of SAJOB on TURN. The positive BETA indicates that as the level of

qualification increases so does the slope of the SAJOB - TURN regression line. This is illustrated when the sample is broken into qualification groups (Table 7.4). The strength of the positive correlation is generally greater at higher levels of secondary qualification. Despite a relatively small group size the School Certificate and Bursary grouping have significant SAJOB - TURN correlations whereas the No Qualification group does not.

Table 7.4

Correlations Between SAJOB and TURN for Five Levels of Secondary School Qualification

	Qualification	Number	Percent	r	Prob
1	No Qualification	n = 236	63.3%	-.027	p>.324
2	School Certificate	n = 78	20.9%	.225	p<.024
3	University Entrance	n = 43	11.5%	.151	p>.167
4	Higher School Cert'	n = 5	1.4%	.159	p>.399
5	Bursary	n = 11	2.9%	.601	p<.025

Apart from the impact of SFE on TURN showing that those rated higher on application tended to be stayers as did older workers, a number of significant main effects were observed. Gender was positively related to TURN showing that females were more likely to leave before the end of the season. The main effect of age showed that older workers were more likely than younger workers to complete their seasonal contracts.

Self-assessment and absence. The impact of SFE only on AB-DURATION was significant (Adjusted $R^2 = .01948$; $F = 8.4290$, $df\ 373, 1$, $p < .004$). The entry of SAJOB did not produce a significant R^2 change. With all variables entered into the analysis the overall R was $.22565$ (Adjusted $R^2 = .02484$; $F = 1.9528$, $df\ 10, 364$, $p < .04$).

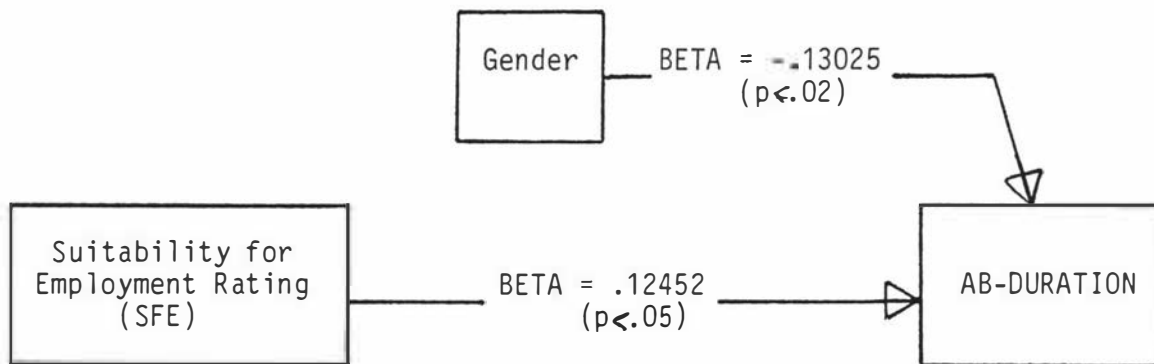


Figure 7.5. Moderated Regression Results for AB-DURATION (length of time absent from work without a telephoned reason).

As shown in Figure 7.5, no moderator effects were significant. SFE was significantly related to AB-DURATION suggesting that applicants receiving higher SFE ratings were more likely than lower rated workers to have more time off work for no given reason. The negative relationship of gender (male coded 1, females coded 0) to AB-DURATION indicates that females are more likely than males to have more time off work without a telephoned reason.

The impact of SFE only on AB-FREQUENCY was significant (Adjusted $R^2 = .01179$; $F = 5.4627$, $df\ 1, 373$, $p < .02$). The entry of SAJOB resulted in a significant R^2 change ($.01632$, $p < .02$). At the final stage of the analysis the overall multiple R was $.24826$ ($p < .01$). As can be seen in Figure 7.6, no moderating effects were significant. SFE

did not demonstrate a significant main effect in the presence of the other variables but SAJOB did have a positive impact on AB-FREQUENCY. The direction of the relationship suggests that those rating themselves higher were more likely to have frequent absences during their seasonal contract than those who rated themselves lower. Age was also negatively related to AB-FREQUENCY implying that younger workers were more likely to have more frequent absences than older workers.

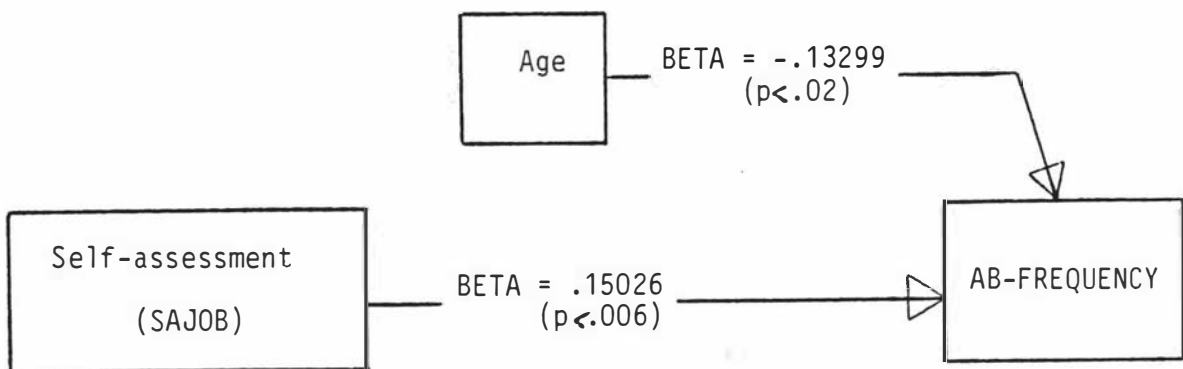


Figure 7.6. Moderated Regression Results for AB-FREQUENCY (number of times absent from work without a telephoned reason).

The effect of job preference on self-assessment. The correlation between job preference (PREFERENCE) and SAJOB was .18 ($n = 375$, $p < .001$, 2 tailed) indicating a significant level of common variance. To evaluate the effect of this shared variance on the predictive relationship between SAJOB and the criteria measures, four stepwise multiple regression analyses were conducted, one for each of the criteria used in this study. SAJOB was stepped in first followed by PREFERENCE. The R^2 change resulting from the entry of PREFERENCE in the analyses was nonsignificant for all four dependant variables (PERFORM R^2 change = .00244, AB-DURATION R^2 change = .00865, AB-FREQUENCY R^2 change = .00424, and TURN R^2 change = .00061).

REFERENCE did not add any additional predictive variance above that offered by SAJOB for the particular criteria used in this setting.

Discussion

The results reported for the predictive utility of self-assessment in Study B were similar to those found in Study A. Relationships were either not significant or in the opposite direction to that expected. The predictive utility offered by self-assessment (SAJOB) above that accounted for by current selection methods was not significant in the case of appraised performance (PERFORM), turnover (TURN), and absence duration (AB-DURATION) criteria. SAJOB was significantly related to absence frequency measures (AB-FREQUENCY) but in the reverse direction to that expected. Applicants giving themselves higher self-assessments were more likely to be absent than applicants giving themselves lower self-assessments. The unique variance offered by suitability for employment ratings (SFE) was a significant predictor for AB-DURATION but not for AB-FREQUENCY. The SFE - AB-DURATION relationship suggested that applicants rated high on employment were likely to spend more time away from work than employees receiving low SFE ratings. SFE ratings were also negatively related to appraised on-the-job performance. High suitability for employment ratings tended to be associated with seasonal workers who completed seasonal employment contracts rather than workers classified as dysfunctional turnovers.

As discussed in relation to Study A, Bandura (1982b) felt that a lack of task and situational knowledge, as well as factors in the selection situation acting as disincentives to accurate self-assessment, could affect the strength of self-assessment - work criteria relationships. Inflation or distortion of self-assessments are not uncommon in selection situations (e.g., Anderson et al., 1984; Ash, 1980; Elliot, 1981;

Farh & Werbel, 1986). It may have been that the temporary nature of this work further exaggerated any such effects.

Over 55% of applicants reported knowing about the work through being told by others. It could be speculated that "local knowledge" regarding the unskilled nature of the work offered to seasonal workers, as well as its casual nature, may have encouraged applicants to feel that they could easily perform well in this type of job. Almost 90% of applicants were employed as belt or general line workers, categories that had the highest average self-assessments of the five work types (7.3 and 6.7 respectively) especially when compared to the driving category with an average of 3.2. Such comparatively easy work may have been seen as simple to perform. What self-assessors may not have considered is that the repetitive work could be difficult to perform over a long period of time. Reactions to highly structured, repetitive, and closely supervised work may have been reflected in appraised performance, as well as turnover and absenteeism criteria (Mowday, Porter, & Steers, 1982), so reducing their relationship with self-assessment.

The tendency for older workers to be more accurate predictors of performance appraisal and AB-DURATION criteria demonstrated in Study A was not supported in Study B. Main effects in the regression analyses showed older workers were more likely to receive high job performance ratings (Figure 7.3) and stay until the end of the season (Table 7.4). The only variable to demonstrate a significant moderating effect was the level of secondary educational qualifications on the SAJOB - TURN relationship (Figure 7.4). Other researchers have suggested educational level may moderate the relationship between a self-assessment and some achievement measure in non-work settings (e.g., Bailey & Bailey, 1971; Hansford & Hattie, 1982; Kooker, 1974).

Other studies conducted in work settings have found no such effect (e.g., Brief et al., 1977; DeNisi & Shaw, 1977).

In general terms, seasonal applicant sub-groups having any secondary school qualification had a positive SAJOB - TURN association (Table 7.4). Applicants having School Certificate (sat in the third year of secondary school) and Bursary (sat in the fifth year) displayed correlations of .225 ($p < .03$) and .601 ($p < .03$) respectively. While it seems possible that those obtaining such qualifications were more capable, therefore being able to exercise a better perception of ability-related performance (Mabe & West, 1982), such an effect was not linear and did not appear in Study A.

The possibility that differential access to job information would have some moderating effect on the accuracy of self-assessment in both studies was not supported in either investigation. This may be because the work content is relatively easy to understand no matter what the source of information. This is particularly so when compared to a managerial or technical position even though such aspects as the closely supervised and highly repetitive nature of the work may not be adequately conveyed by any source other than enactive experience.

Gender as a moderator of self-assessment - criterion relationships has been both supported (Dicken, 1969) and discounted (Ekpo-Ufot, 1979; Shore & Thornton, 1986). Gender did not demonstrate a significant moderator effect on any of the SAJOB - work criteria relationships in Study B but was directly related to TURN and AB-DURATION. Males were more likely to be classified as stayers and have less time absent from work than females. No such main effects were demonstrated for gender in Study A.

While not the focus of this study, the SFE rating acted as the point of comparison for self-assessment. The basis of the SFE rating may be responsible for its poor predictive utility. These ratings were based on a brief meeting between a personnel staff member and an applicant that was structured in terms of the biographical information being recorded but did not involve any job related questions beyond establishing availability for work. Where applicants had a history of seasonal work with the organisation, past work performance (if recorded) was taken into account. For those applicants with no recorded work history, judgments of job suitability appear to have been based on very limited information.

The nature of the dependent variables used in Study B are partly reflected in the level of common variance shared by some. Firstly, the relationship between PERFORM and measures of both turnover and absenteeism were strong and in the logically expected direction. Higher performers tended to be stayers, lower performers dysfunctional turnovers. Higher performers tended to have lower rates of absenteeism whether measured on the basis of time away from work or the number of absences. It was reported by personnel and supervisory staff that absences and withdrawals were disruptive. The significant relationship found between PERFORM and TURN did not occur in Study A. There was no relationship between TURN and the two measures of absenteeism. This lack of association may be due to the different consequences of turnover and absenteeism, the spontaneity of absence versus turnover decisions, or the possibility that one action substitutes for the other (Mowday, et al., 1982; Porter & Steers, 1973).

Of peripheral interest was the relationship between AB-DURATION and AB-FREQUENCY. The relationship between the different measures of absenteeism is

stronger than that reported for permanent female workers (Nicholson & Goodge, 1976).

The type of scale used in Study B was a vertical graduated line with the statements "One of the Best" at the top and "Not Too Good" at the bottom (Appendix 7). As in Study A, self-assessments were skewed in a positive direction although 16% of ratings were at or below the midpoint of the scale. The nature of the scale as well as the verbal and pictorial job descriptions may have influenced the level of variability. During the first three days of worker registration, no applicants had or reported trouble understanding the self-assessment requests. Whether or not the changes to the self-assessment format were in any way responsible for the change in the self-assessment - job performance relationship is unknown.

In Study A and Study B self-assessment did not contribute to the prediction of job performance. Inconsistent moderator effects across both studies in this setting means no one factor can offer an explanation for the poor predictive validity observed. The seasonal nature of the work contract may have encouraged applicants to treat the self-assessment request in a less serious manner than if they had been applying for a permanent position. Unfortunately no equivalent sample applying for permanent work was available to test this possibility.

CHAPTER EIGHT

EVALUATION OF SELF-ASSESSMENT VALIDITY IN A POLYTECHNIC NURSING COURSE

The Polytechnic Nursing and Health Studies Department was responsible for the tuition of nurse trainees to the stage where the trainees were granted permission to sit an examination for registration as a comprehensive nurse. The course used for evaluation was one of 15 conducted in similar settings under the control of the Department of Education. It covered components of theoretical instruction related to clinical practice over a period of 3 years. Minimum requirements were set by the registering authority, the Nursing Council of New Zealand.

This particular nursing course was unique amongst organisations taking part in this study in that it required applicants, as well as nominated referees, to complete a reference form. The form was developed by the organisation and had been in use for two years prior to this study. An initial request to the Polytechnic resulted in access being given to the self-completed reference forms. The forms involved a series of open-ended questions and were identical to those sent to referees. The purpose of Study A, the first of two in this setting, was to examine the relationship between a self-completed reference form and success in the first year of the training programme and to make recommendations on the design of a self-completed reference form for use with subsequent student intakes.

Study A

Method

Subjects

Subjects were 129 student nurses admitted to a comprehensive nurse training programme. Ages ranged from 16 to 44 years, an average of 19 years and four months, with 84% of subjects falling into the 16 to 20 year age range. Four were male and 125 were female. Forty six (36%) had not passed the Sixth Form University Entrance qualification, 83 (64%) had passed University Entrance. One trainee had completed a university undergraduate degree.

Procedure

Self-assessment requests. The source of self-assessment information was a reference form developed by nursing course staff which asked applicants to comment on listed attributes such as integrity, honesty, acceptance of responsibility, health, personal appearance, relationships with other people, and attitudes to assigned tasks or work (see Appendix 8). The form also asked whether or not the applicant was suitable for a nursing career and able to cope with the academic demands of the comprehensive nursing course. Requests for comment were open-ended such as "Do you think he/she has the ability to cope with the academic demands of a 3 year full time course?" The format had been developed by the organisation based on the experience and resources of staff members. Although it was not clear whether any pilot testing of the form had been done, staff responsible for selection of candidates felt applicants clearly understood the form.

Criterion. According to nursing course staff, most students who failed the course did so in the first year. Failures in subsequent years were extremely rare. At the end of year one, student course performance was evaluated by nursing course

tutors based on the end of year exam, performance in the clinical setting, and course performance during the year. Students were coded as a pass ($n = 109$, 84.5%) or fail ($n = 20$, 15.5%) by all course tutors. Exam marks or clinical performance reports were not available as criteria.

Selection procedures. Each Polytechnic nursing course was expected to provide places for suitably qualified applicants in the immediate geographic area although course numbers could be made up to strength by taking candidates from anywhere in New Zealand. In this study the Polytechnic did not use interviews for selection but instead relied on information in the application form, educational qualifications, and self- and referee-completed reference form information. The first face to face contact most applicants had with Polytechnic staff was in group "information giving sessions" in the first week of the course. Some students had visited the Polytechnic to obtain course information, application forms, etc., while others had met tutors giving career information talks at secondary schools. These contacts were not a part of the selection procedure.

Students completed an application form which required biographical, health, and educationally related information and the self-completed reference form. On receipt of the completed forms the Polytechnic contacted referees and asked for the reference forms to be completed. The applicant pool was sorted by Polytechnic staff into acceptable, marginal, and unacceptable groups based on application form and completed reference forms. Specific criteria were not defined. As stated by one staff member, "how do you compare a 17 year old person with university entrance and a 35 year old with older children, no secondary school qualifications, but with experience in caring for others?"

Coding of self-completed reference forms. Polytechnic staff involved in selection evaluated the self-completed reference form as to how positively applicants were described and how consistent they were with referee-completed references. To code the open-ended answers each item was rated on how positively the applicant was described using the following scale.

1. Exceptionally Positive. (excellent, exceptional, etc.)
2. Positive (without any qualification)
3. Neutral or Positive Statement with Qualification (e.g., Yes)
4. Indication of Negative Aspect
5. Negative Statement

Ratings of all forms were made by two graduate psychology students. Unfortunately no Polytechnic staff responsible for selection were available for the rating process although they agreed that the rating scale itself was a good representation of how they themselves treated the responses. Rater training involved an explanation of the rating scale and how they were used, practice using actual items, and discussion of these practice items. Each rater then rated a set of 20 forms which were not part of the study sample. The ratings were then compared and discrepancies discussed. At this point both raters felt they were rating the items in the same way. Each rater then did the rating task on the sample of 129.

Positiveness ratings were made on each item. Forms were missing from student files in some cases. These had been removed by the nursing department as the information they contained was regarded as too personal to be retained on file. In addition to rating individual items, an overall assessment was made using the same rating scale.

Results

Inter-rater reliability. Correlations between raters were computed to give an indication of the level of inter-rater agreement (Table 8.1). Only items with correlations above .5 were retained. This level is regarded as satisfactory for further analysis in this type of research (Baxter, Brock, Hill, & Rozelle, 1981). Adoption of this cutoff point resulted in the rejection of nine items and the rating of overall positiveness. The items that were retained for further analysis are listed in Table 8.1. The judges' ratings were averaged for the purposes of further analysis.

Table 8.1

Inter-rater Reliability Coefficients

Coded Open-ended Items	Coefficients
Integrity	.5087*
Honesty	.3054
Acceptance of Responsibility	.2980
Health	.5875*
Personal Appearance	.3342
Special Abilities	.5066*
Relations with Persons of Own Age	.5970*
Relations with Persons in Authority	.5713*
Acceptance of Correction/Development	.4007
Cooperation	.3551
Perseverance	.5611*
Concerns for Needs of Others	.3707
Suitability for Nursing Career	.1292
Academic Ability	.4431
Overall Self-completed Reference Form	.4199

* items retained for further analysis

Predictive validity of self-completed reference form information. As the criterion involved two categories (pass and fail), a linear discriminant function analysis (LDF) was used to evaluate the predictive validity offered by that information surviving the inter-rater reliability check. Whereas regression analysis is designed to

predict a position on a criterion scale, LDF is used to classify cases into groups (see Chapter Six).

A Wilks backwards stepwise LDF was performed using the Statistical Package for the Social Sciences (SPSS Inc, 1986). A backward rather than a forward stepwise method was preferred in line with Hull and Nie's (1981) recommendation that it is the more economical method and is appropriate when it is felt most variables will contribute significantly to discrimination. In addition, it allows the contribution of all items to be assessed before an optimal, more economical solution is arrived at.

None of the 6 coded items were significantly related to success in year one of the nursing course and therefore did not qualify for entry into the LDF.

Discussion

The level of reliability demonstrated by coding the information was barely acceptable. Although the .5 cutoff level adopted for acceptance of variables for further analysis has a precedent in similar research (Baxter et al., 1981; Meyer, 1970), this means only 25% of inter-scorer variance was accounted for. Even at this level many items were lost for further analysis. Open-ended formats should not be neglected for the rich insightful information they produce, as well as encouraging respondents to express themselves in their own language (Sudman & Bradburn, 1983). Carefully designed pre-coded formats, however, avoid judgments such as those required in the present study. If such qualitative information is required, it would seem that a combination of formats could be considered. The level of reliability demonstrated indicates that findings based on coded information are at best only tentative.

The absence of any relationship between self-assessment items and year one course success could be due to the characteristics of the request for self-evaluative information. Mabe and West (1982) showed that factors such as a match between self-assessment requests and criterion, expectations of validation, and requests made relative to a comparison group all enhance the validity of self-assessment requests.

Firstly, predictor and criterion dimensions were not matched. Student assessment was based on professional standards, clinical competence, and theoretical knowledge. Self-evaluation items passing the reliability check (Table 8.1) were only indirectly related to the criterion of first year success. The "relations with persons in authority" may have some bearing on the assessment but "academic ability", which referred to the ability to cope with the academic demands of the nursing course and could be considered the most directly related, did not survive the reliability check. It is possible that the lack of match between items acting as predictors and the first year success criterion contributed to poor validity. Secondly, the request for a self-completed reference form did not contain any suggestion that the information would be checked or validated so increasing the desire to give accurate rather than enhanced information. Thirdly, the request for self-completed reference data did not mention a comparison group although it would seem obvious that applicants knew they would be compared with other nurse trainee applicants.

As a result of this first study a number of recommendations for changes in the self-completed reference form were made to the Nursing Department. These were:

1. That pre-coded formats should be used in the self-completed reference form where possible and in conjunction with open-ended questions.

2. That self-assessments should be based on how performance in the course is judged.
3. That a description of the characteristics of successful and unsuccessful nurse students be included with a self-assessment request to enable comparative judgments to be made by applicants.
4. That self-assessment instructions include a request for accuracy accompanied by an explanation that the assessment would be compared to course performance.

Study B

The recommendations arising from Study A were aimed at providing information for use in the design of a future self- and referee-completed reference form. However, an unexpected initiative at a conference of nursing course executives resulted in a decision to standardise the format of reference requests for all Polytechnic nursing courses. The recommendations made by the first study were taken into account but the final design was also influenced by the requirements and preferences of other Polytechnic nursing departments.

Study B also provided the opportunity to evaluate a number of potential moderating effects. Three individual differences focused on in this setting were intelligence, locus of control, and social desirability. These moderators were included for the following reasons.

- a) As discussed in Chapter Four, it has been suggested that more intelligent subjects tend to be more accurate self-assessors (Mabe & West, 1982).

- b) Research indicates that internal locus of control is associated with accurate self-evaluation of ability (e.g., Steger, Simmons, & Lavelle, 1973).
- c) It seems that the tendency for individuals to describe themselves in socially desirable terms may affect the accuracy with which these individuals make self-assessments (e.g., Jones & Pittman, 1982).

Method

Subjects

The subjects were 90 comprehensive nurses who completed a three year Polytechnic training course. A further six cases were not used in the investigation as they had withdrawn from the course for reasons unrelated to course performance (family leaving area and long term illness). Ages ranged from 16 to 44 years, an average of 19 years and six months, with 63% of subjects falling into the 17 to 18 year age range. Five were male, 85 were female. Seventy two (80%) had passed the Sixth Form University Entrance qualification. Fifteen trainees had done some university undergraduate study. Eighty one cases were classified as European and nine as Polynesian (Maori, Pacific Islander, Asian). Exactly half of the sample had previously worked in a nursing job.

Procedure

Self-assessment requests. The self-completed reference form was identical to that used by referees (Appendix 9). The form consisted of a number of sections. Firstly, six semantic differential scales required ratings of personal qualities that were regarded as a primary requisite for success in nursing (honesty, maturity, reliability, grooming, tolerance, and acceptance of responsibility). Secondly, open-ended comments were requested on interpersonal relationships as well as attitudes to work

and study. Thirdly, candidates were asked to respond yes or no to questions such as "do you consider the applicant able to undertake a full time 3 year course of intensive study?".

Criterion. At the end of the third year of the comprehensive nursing course each student was assessed. This final year assessment (FINAL) is based on tutor evaluation of professional attitudes and responsibilities, clinical competence assessed by tutors in work settings, and theoretical knowledge assessed by examinations. Students may receive an A, B, or C pass, repeat the year, or fail the course. These categories were coded 4, 3, 2, 1, and 0, respectively. Categorisation of students occurs only after a number of assessment meetings where all tutors involved in student training discuss marks and grades given for tests, assignments, examinations and clinical work. The final assessment was used as a criterion in preference to year one success. Polytechnic staff felt that second and third year dropout rates had increased following Study A and as a result year one success was no longer a good indicator of overall course success.

Selection procedure. The selection procedures were the same as described in the first study conducted in this setting. The course intake had been increased from 60 to 100 in this particular year in anticipation of an increased demand for trained nurses. Only 96 places were filled in this particular intake due to a shortage of suitable applicants.

No attempt was made to code the open-ended responses due to the poor reliability demonstrated in Study A. Responses to the scaled items were coded five high and one low. For yes/no items "no" was coded zero and "yes" was coded one.

Moderator variables. Eight potential sources of moderator effect were examined in this study. Biographical data contained in the application form provided information on race (European/non-European), age (years and months), gender (male/female), education (number of UE subjects passed) and work experience in nursing (yes/no). During the first six weeks of the comprehensive nursing course the researcher ran a series of workshop sessions with students. During these sessions students were asked and agreed to complete three additional measures. These measures were of intelligence (Raven, 1965), locus of control (Rotter, 1966), and social desirability (Crowne & Marlowe, 1964).

Intelligence was measured using the Raven Advanced Progressive Matrices (Raven, 1965). This test has been extensively evaluated in a wide variety of situations (Buros, 1965) and is regarded as particularly useful in assisting selection of trainees for advanced technical studies (Raven, 1965).

Rotter's (1966) scale was preferred as a measure of locus of control in that it is the most widely used scale of this type, it is generally unrelated to measures of social desirability (Robinson & Shaver, 1973), and it has been used in investigations involving self-estimates of individual abilities (e.g., Minor & Roberts, 1984).

Social desirability was measured using Crowne and Marlowe's (1964) Social Desirability Scale (SDS) as a common measure of this construct, suitable for a normal population.

Results

Predictive validity of self-assessment. In response to the questions, "do you consider 'the applicant' able to undertake a full time 3 year course of intensive study?" none of the sample responded "no" thus demonstrating no predictive utility for the question.

The other items relevant to course and nursing performance were the six personal characteristics (honesty, maturity, reliability, grooming, tolerance, and ability to accept responsibility). Possession of these personal qualities were regarded by the Polytechnic as primary requisites for success in nursing (Appendix 9). As it was possible that the six items were being answered in the same way despite the fact that they were requesting ratings on different personal aspects, a test for homogeneity of the six items as a scale was conducted. Cronbach's alpha was .68, indicating a moderate level of homogeneity. A further check was made by examining the item to total correlations which essentially is the relationship each scale has with the other scales. The correlations ranged from .33 to .54 with an average of .43. As these relationships were not strong it was decided to analyse each item separately.

The range of ratings on all personal quality items was restricted to a medium rating (3) and above. In addition, medium ratings comprised no more than 10% of the ratings made on any one of these items. The average ratings were, honesty 4.89 (SD = .32), maturity 4.30 (SD = .55), reliability 4.71 (SD = .50), grooming 4.42 (SD = .62), tolerance 4.36 (SD = .66), and acceptance of responsibility 4.60 (SD = .56).

A moderated regression analysis was conducted for each of the six personal characteristic items. The item was stepped into the analysis first followed by the variables being considered as moderators to partial out the main effects. The last

step in the analysis was the entry of the variables which were the products of the potential moderators and the personal characteristic rating. Not one of the self-rated personal characteristics was significantly related to FINAL; nor was the R^2 value on the final step in each of the six moderated regression analyses significant (Table 8.2).

Table 8.2

R^2 Values on First and Last Step of Moderated Regression Analysis

Characteristic	R^2 Self-Assessed Rating	R^2 Total Analysis
Honesty*	$R^2 = .0018$; $F = .1547$, df 1, 88 NS	$R^2 = .1691$; $F = .9288$, df 16, 73 NS
Maturity	$R^2 = .0079$; $F = .7043$, df 1, 88 NS	$R^2 = .2107$; $F = 1.1308$, df 17, 72 NS
Reliability	$R^2 = .0047$; $F = .4146$, df 1, 88 NS	$R^2 = .2421$; $F = 1.3531$, df 17, 72 NS
Grooming	$R^2 = .0066$; $F = .5877$, df 1, 88 NS	$R^2 = .2383$; $F = 1.3250$, df 17, 72 NS
Tolerance	$R^2 = .0136$; $F = 1.2122$, df 1, 88 NS	$R^2 = .2887$; $F = 1.7190$, df 17, 72 NS
Acc' Res'	$R^2 = .0002$; $F = .0179$, df 1, 88 NS	$R^2 = .1698$; $F = .8663$, df 17, 72 NS

df values for total analysis differ from other analyses as one moderator variable did not meet tolerance levels.

With the exception of the analysis involving **grooming** as a personal characteristic, no analyses revealed significant moderator effects. Of the eight potential moderators, passing or not passing the sixth form university entrance exam (UE) was the only one which showed a significant effect ($BETA = .36242$, $p < .04$). The positive BETA suggests that cases with more UE subjects had a steeper sloped regression line than those with fewer UE subjects. At this final step of the moderated regression analysis, grooming demonstrated a low negative relationship with FINAL for the whole sample ($BETA = -$

.02117, NS). grooming/FINAL correlations calculated separately for those with zero, two, three, four, and five UE subjects showed that cases with four or less subjects had negative Grooming/FINAL correlations, whereas those with five subjects had a weak but positive correlation. No relationships were significant and more importantly, were not shown to exist for other self-rated personal characteristics.

In the presence of each of the self-rated personal characteristics locus of control (LC) had a significant negative relationship with the FINAL criterion. The correlation of LC with FINAL was $-.21$ ($p < .03$) signifying that cases tending to a belief in an external locus of control were more likely to get a lower score on FINAL.

Conversely, cases who have an internal expectancy were more likely to receive higher scores. No other moderators demonstrated strong or consistent relationships with FINAL across any other analyses.

Discussion

Study B showed self-assessment information requested using a self-completed reference form offered no predictive utility above that of the current selection procedures. No consistent indication of a moderating effect was found. The weak and isolated moderating effect shown by UE passes on the grooming/FINAL relationship provides little indication of other than a chance effect. Failure to detect other potential moderators may have been due in part to the low variability in self-assessment ratings. The restricted range of responses was demonstrated in "yes only" answers to the question "do you consider 'the applicant' able to undertake a full time 3 year course of intensive study?". Although self-assessments often demonstrate less variability than assessments by others (Meyers, 1980, Thornton, 1980), in this

particular case it would seem illogical for an applicant to state that he or she would not be able to cope with the course being applied for.

The restricted range of responses was also apparent in responses to self-ratings of personal characteristics. It is understandable that applicants would rate themselves high on characteristics that are described as primary requisites for success in nursing. It is also possible that the lack of variation in range could have been due to the preselected nature of the applicant sample.

The character of the self-assessment requests may have affected their relationship with course performance. What was actually meant by the various personal characteristics dimensions may not have been clear to applicants who could have had different ideas as to what the various personal dimensions meant. The ability to tolerate an abusive patient may be quite different to the ability to tolerate other students talking in a school library. Ratings may therefore have varied as a function of applicant's conception of their meaning (Primoff, 1980).

It may have been more appropriate to ask applicants to state the level at which he or she would expect to perform in a group of comprehensive nurse trainees in specific (and defined) areas. Suitable comparison group information for specific items or dimensions could also be supplied to applicants. While the self-completed reference form made no mention of validation of the ratings provided by applicants, it is difficult to see how the predictive accuracy of rating can be improved when the basis of the required ratings are unclear.

Apart from demonstrating a lack of predictive validity for self-assessment information in this form and an absence of consistent moderator effects, Study A and Study B

provided useful qualitative information. Firstly, the concept of a self-completed reference form is an interesting approach by a New Zealand organisation and would seem worthy of further investigation. Polytechnic staff indicated that they introduced the process because it gave applicants an opportunity to provide some unique information which would otherwise be unavailable to those responsible for selection. The quality of this information could be improved, however. For example, including suitable information with a self-assessment request would allow all candidates to make ratings based on the same information. A forced-choice format also could be incorporated (Carroll & Nash, 1972).

Secondly, the experience of this setting showed that despite demonstrating the difficulty of obtaining an acceptable degree of reliability, administrators responsible for selection persisted with the use of open-ended response items. While the reasons for this are not clear, reluctance to quantify predictor information may have been due to a number of factors, including the political context in which the form was designed.

Thirdly, practitioners appear to prefer the use of titled but undefined personal characteristics as descriptive dimensions in selection settings. It can be implied that they expect all applicants to have a similar understanding of those dimensions, an understanding which is similar to their own. The fallacy of this approach is not apparent to those concerned.

Interestingly, as part of the process of monitoring course performance, nurse trainees were required to complete self-ratings of work completed in various modules of the course. Tutors have said that at first students found the process difficult but as more self-assessments were made this apprehension disappeared. Also as students

gained more self-assessment experience they appeared to become more accurate (as far as the tutors were concerned) and treated the process as constructive. They were willing and able to justify self-ratings based on their own perceptions rather than deferring to what they believed a tutor's opinion to be.

Following this investigation, the organisation continued to use self-completed reference forms. More information on the nursing course was to be sent with the forms to applicants and referees. No statistical analysis of the content of the self-completed or referee-completed forms was to be done. Admission decisions were to be made mainly on the basis of educational qualifications and the application form, although the responses on both types of reference form are also to be considered.

CHAPTER NINE

EVALUATION OF SELF-ASSESSMENT VALIDITY IN A GOVERNMENT DEPARTMENT

Many New Zealand graduates apply every year for positions in the Public Service. A recent research project on graduate work preferences conducted by George and Hyde (1984) gave some indication of how graduate selection is conducted in government departments. Information from over 100 graduates showed that selection was almost always based on various forms of selection interview. Exceptions included the use of psychological assessment and work sample by one other department in the Public Service (Shouksmith & Petersen, 1984).

The particular department approached for this study was an employer of a large number of graduates and used selection methods other than those involving just an interview. Although having a preference for various commerce and economics qualifications, the department employed graduates with a range of specialties. The investigation was conducted as two studies because selection procedures changed considerably halfway through the project.

Study A

Method

Subjects

The subjects were 28 successful applicants for assistant advisory officer positions. These applicants were from a pool of 54 who had been pre-selected by screening interviews from a pool of approximately 240 applicants and invited to attend

organisation selection procedures. Twelve cases were male, 16 were female. The average age was 24.8 years with a range of 21 to 37 years. Seven cases had completed or were in their final year of a Bachelors degree and 21 cases had completed or were in the final stages of a post-graduate degree (e.g., Masters degree).

Procedure

Development of self-assessment requests. The work content of the positions to be occupied by recruited officers could not be defined prior to selection. Newly appointed officers could be placed in any one of seven divisions dependent on changing divisional requirements and where possible, an officer's preference. In addition, the role of the department was changing from an emphasis on the enforcement of government regulations to promotion of trade and commerce. The job content of the assistant advisory officer position was expected to change but there was no clear indication as to what changes were involved. As it was not possible to request self-assessments of work in specific areas, applicants were asked for self-evaluations of overall job performance in the assistant advisory officer position.

One aspect common to all jobs was the way in which the organisation judged job performance 12 months after appointment. The Public Service performance review system required judgment on dimensions of work such as organisation of work, output, communication, versatility, initiative, and staff relations. It was decided to include the same performance review items in the request for self-assessment as they were "matched" to later evaluations and were standard across all positions. The specific items used can be seen in Appendix 10.

In addition to self-assessments of work dimensions, applicants were asked to make two assessments of their future overall job performance. One overall job performance item used a similar five point format to that used for promotion rating in the department. The item asked the candidate how effective they thought they would be in fulfilling the major duties, responsibilities, and objectives of the job. The second overall performance assessment asked candidates to rate themselves on how well they would do the job based on their performance over the assessment day. This item was intended to be equivalent to the rating given candidates following the interview and used the same A+ to C- scale.

A draft self-assessment form was developed containing the items described and discussed with departmental management. With the exception of the self-rating of job performance based on the assessment day, the options in all items were derived from department material. The form was pilot tested on eight department officers employed within the previous six months and on a group of 12 graduate students. The pilot group was asked to complete the form as if applying for a job and to note anything that was unclear or wrong. No changes to the form were made as a result of the pilot test.

Criteria. At the time of this study the New Zealand Public Service was introducing a new performance review system and organisational delays in this setting resulted in few officers having their performance appraised. As this meant official criterion information was not available for most of the sample, an alternative measure had to be used. Approximately every 12 months all employees in the organisation were graded and awarded salary increases. Grades were based on an officer's job performance as viewed by divisional management. An officer could be awarded three, two, one, or no salary steps, or rated as an unsatisfactory performer. These outcomes

were coded 5, 4, 3, 2, and 1 respectively and were used to form the criterion PROMOTE.

Promotion data were not available for all employed officers as five officers had resigned within six months of employment and managers were not prepared to provide promotion data for this group.

Eight of the employed applicants resigned within the first 12 months of employment. As this number was a sizeable proportion of the sample (29%), subjects completing their first year of employment were coded one and those resigning beforehand were coded zero to form the criterion TENURE.

Selection procedures. The 54 applicants who attended selection procedures had been chosen from an applicant pool of around 240 on the basis of a screening interview and application form information. The 54 applicants attended an "assessment day" in groups of eight to 10 held at the organisation's head office. The selection procedures were made up of the following components.

- a. **Introductory Session:** A formal presentation by two middle managers detailing the broad functions of the department, training opportunities, conditions of employment, salaries, and career paths.
- b. **Report Writing Exercise:** Applicants had to evaluate a letter from a member of the public and provide a report including a recommendation of action to be taken.
- c. **Meeting with Incumbents:** Applicants were introduced to officers currently employed who had been asked to show the candidates around the work place and answer any enquiries candidly.

- d. Panel Interview: Applicants were individually interviewed by a panel of three middle managers.

The interviewers came to an agreement on a candidate's job suitability on the basis of the panel interview in most cases. The written exercise was not marked in any standard way and only referred to if the panel felt further information was necessary. The decision of the panel was then discussed with the chief executive of the organisation and a final decision on an offer of employment was made. At the last stage of the assessment day, applicants were asked to complete a self-assessment.

Results

Predictive validity of self-assessment: A complete set of ratings made by the interviewing panel was not made available so could not be incorporated in the analysis. The sample size was not sufficient for multivariate analysis and therefore correlation coefficients were used. Table 9.1 shows the correlations between self-assessments and the criteria PROMOTE and TENURE.

Two subjects did not complete a self-rating of future performance based on the assessment day. When questioned, both stated that they had received no feedback on their performance so they did not have enough information on which to make a rating. Interestingly both these subjects made self-ratings of expected future performance. None of the relationships between self-assessment and work performance criteria were significant although the difference in direction of the relationships is of some interest. A self-assessment of future job performance is positively related to PROMOTE but when such a rating is based on performance over

the assessment day the relationship becomes negative. The relationship between the two self-ratings was 0.38 ($p < .05$) and between TENURE and PROMOTE 0.39 ($p < .05$).

Table 9.1

Predictive Relationships Demonstrated Between Self-Assessment and Work Criteria

<u>Self-Assessment</u>	<u>Work Criteria</u>	
	PROMOTE	TENURE
Future Job Performance	0.31 (23)	-0.10 (28)
Future Job Performance (based on assessment day performance)	-0.15 (22)	0.18 (26)

Self-assessed dimensions of work and work criteria. The relationships between the 11 dimensions of work on which self-assessments were requested and the two work criteria were examined. Two tailed tests of significance were used due to the exploratory nature of the analysis. No self-ratings of dimensions of work were significantly related to TENURE. Oral and written communication ($r = .42$, $p < .05$) and initiative ($r = -.56$, $p < .005$) were the only dimensions significantly related to PROMOTE.

Discussion

The above results show that self-assessments of overall work performance offered little in the way of predictive validity. The low level of relationship demonstrated by

the overall self-assessments was similar for self-assessments of various dimensions of work.

The positive and negative direction of the relationship between self-assessment and the work criteria was of some interest even though the associations were not significant. When candidates were asked to make a self-assessment based on their performance over the assessment day, the assessment was negatively related to job performance and positively related to tenure. The direction was reversed if subjects' self-evaluations were not requested on any specific basis. Whether such a change in direction reflects the information on which the self-assessment requests were made is difficult to say. The relationship between the two self-assessments ($r = .38, p < .05$) was not as strong as would be expected between similar measures of the same performance, as for example, in a test of reliability. It was decided that further investigation in this setting would be informative.

As for the overall self-assessments of work performance, there appeared little consistency in the relationship between assessments of dimensions of work and work performance criteria. The two dimensions that were significantly related to work performance, communication ($r = .42$) and initiative ($r = -.56$), were in opposite directions. That self-perceived good communication skills and a lack of initiative describes those officers receiving high performance appraisals is intriguing. Rather than taking these associations as a reflection of public service promotion processes, however, it should be remembered that there were 20 further correlations in the same analysis that were not significant.

This study suggests that self-assessment is acceptable to applicants and can be easily used within the framework of an existing selection procedure. The lack of a

significant association between self-assessment and work criteria may have been due to moderating effects which could not be investigated with this small sample. A second study was needed in this setting involving a larger sample and better researcher access to applicant and criterion information.

Study B

The organisation agreed to a further study with improved access. In comparison to Study A, selection procedures changed to an assessment centre approach and the department was to increase recruitment numbers to around 50 or more officers. In addition, it became possible to include self-assessment as a first and last step of the selection procedures.

Method

Subjects

The subjects were 85 applicants for assistant advisory officer positions in a government department. The sample had been pre-selected by screening interviews from a pool of applicants applying during a recruitment programme covering all seven New Zealand universities. The total sample participated in the assessment centre selection procedures and a sub-sample of 47 received and accepted offers of positions within the organisation. Of the 38 cases not employed within the department 21 were not offered positions, 16 refused a job offer, and one case accepted an offer to continue university study under a bursary scheme. The characteristics of the samples are shown in Table 9.2.

Table 9.2

Characteristics of the Total Sample and the Employed Sample

Characteristic	All Screened Applicants	Employed Applicants
Gender		
Male	39 (45.9%)	18 (38.3%)
Female	46 (54.1%)	29 (61.7%)
Tertiary Education		
No Post-graduate Degree	47 (55.3%)	25 (53.2%)
Post-graduate Degree	38 (44.7%)	22 (46.8%)
Previous Employment		
No Full Time Job	36 (42.4%)	19 (40.4%)
Had Full Time Job	49 (57.6%)	28 (59.6%)
Government Service		
No Work Experience	54 (63.5%)	29 (61.7%)
Govt. Work Experience	31 (36.5%)	18 (38.3%)
Age		
(range, average)	17-34 yrs, 23.5 yrs	20-32 yrs, 23.8 yrs

As Table 9.2 shows, there is little difference between the employed group and the total sample of assessment centre participants apart from there being more females in the employed group. Applicants had university qualifications in a wide range of disciplines including the social and biological sciences, the humanities, law, and commerce. All applicants refusing offers of employment stated they had alternative positions which were more attractive for a variety of reasons such as pay, career, and training opportunities.

Procedure

Development of self-assessment requests. Development of the assessment centre was based on an extensive job analysis. This involved a study of existing job

descriptions, observation, extensive interviewing of advisory officers and their superiors, a survey of incumbents, and meetings with management. Information was collected on work environment, education, job responsibilities, skills, and abilities needed. Based on this information and critical incident data, a list of job activities was compiled. A survey of job activities involving all entry level staff was conducted to verify job content. Meetings with middle and senior management were held to establish which job activities were important to the functioning of the department, which activities involved on the job training, and which were to form the basis of a selection programme. In conjunction with managers, a series of work samples was developed and pilot tested.

The self-assessment request was based on the content of the assessment centre (see Appendix 11). Self-assessment ratings were requested on seven work sample exercises and overall job performance using the same 10 point rating scale (10 = exceptional performance, 1 = very poor performance) as used by assessors. Self-assessment items included a description of the work to be rated and a numerical rating scale (Appendix 12).

The self-assessment request used before and after the assessment centre was pilot tested using 10 recently employed assistant advisory officers and 15 graduate psychology students. Those participating reported no problems with the instructions and felt that the item descriptions adequately reflected detailed descriptions of the work samples. Only minor grammatical changes were needed as a result of the pilot testing. The pre-centre self-assessment form also asked whether applicants had previous self-assessment experience. The post-centre form requested self-assessment of performance in the actual assessment centre as well as a rating of future job

performance. Self-raters were asked to describe the experience on which they based their self-ratings.

Criteria. Quantitative performance data was not readily available in this organisation. At approximately the same time as this study was conducted the New Zealand Public Service introduced a new performance appraisal system that required a review of performance, individual qualities, characteristics, and qualifications as well as specifications of incumbent training to be undertaken in the following year. Unfortunately the resulting information was qualitative only. In addition, employed officers' performance review information was not released to the researcher by some managers and in other cases the reviews were simply not completed.

To provide quantifiable criterion data the direct superiors of all officers in the sample were individually approached and asked to complete a separate performance evaluation form. This form requested ratings on the same seven work sample tasks used in the assessment centre and a rating of overall job performance based on comments made in the qualitative performance review. Where possible, managers two levels above the officers were also asked to complete performance evaluations to allow a reliability check. Inter-rater reliability was acceptable ($r = .91, p < .001, n = 23$). These ratings were made 12 months after the officers were employed and formed the criterion PERFORM. Performance evaluations were obtained for all but one officer because in that case both the superior and the officer had resigned.

Employed officers in the sample were also individually approached and asked to complete a performance appraisal form similar to that completed by their superiors. It was explained that this self-appraisal was for research purposes only and would not

be communicated to the department. As in the case of the PERFORM criterion, self-appraisals were collected for all but one employed officer.

As for Study A, all employees in the organisation were graded and given promotion steps on the government salary scale where appropriate. Coding was identical and data on the PROMOTE criterion was available for all employed officers. Turnover was not used as a criterion as only four of the 47 employed cases resigned within 12 months of employment.

Moderator variables. The relationships between self and organisational criteria were examined for possible moderating effects of age, gender, intelligence, social desirability, full time work experience, and education. With the exception of intelligence, measurement of these variables was the same as for the studies reported in Chapters Six and Seven. The Watson-Glaser Critical Thinking Appraisal (CTA; Watson & Glaser, 1964) was used as a measure of intelligence in this graduate sample. The CTA appears to be the most frequently used cognitive ability test in individual assessment in the United States of America (Ryan & Sackett, 1987) and measures general intelligence and certain logical reasoning abilities (Crites, quoted in Buros, 1972). It is described as appropriate for selection and research situations (Helmstadter, quoted in Buros, 1972) and has been used extensively in research and practice (Mitchell, 1985).

Selection procedures. Applicants who were offered an opportunity to participate in selection procedures attended a one and a half day assessment centre at the organisation's head office. The assessment components of the centre consisted of seven work sample exercises, a leaderless group discussion, and a panel interview. Self-assessment ratings, a cognitive ability test, and the Social Desirability Scale

(SDS; Crowne & Marlowe, 1968) were also required as part of normal selection procedures but this information was not used by the organisation to make job offer decisions. Assessors were generally first line managers and panel interviews were conducted by middle managers.

Each candidate was assessed on an exercise by exercise basis. Following the centre, assessors met to integrate and discuss exercise results and make overall ratings of candidates. These ratings were then reported to an executive committee who made the final decisions regarding job offers. Job offers were made at two levels of salary. An "A" offer was for permanent employment at the maximum salary payable to that particular applicant (dependent on type and level of qualification). A "B" offer was for permanent employment at the second level of salary payable. Some candidates were offered a bursary to continue their university studies. This offer was accompanied by an A or B permanent employment recommendation. The employment offer (OFFER) represented the organisational selection decision. An A offer was coded three, a B offer was coded two, and rejections were coded one.

Prior to each assessment centre, candidates were asked to complete a self-assessment form. At the conclusion of the assessment centre the candidates were again asked to complete an equivalent self-assessment form.

Results

Predictive validity of self-assessment. Two regression analyses were conducted to assess the addition to predictive validity offered by self-assessment over current selection procedures for the performance appraisal criteria (PERFORM) and the promotion criteria (PROMOTE). The employment offer (OFFER) was entered into the

analysis first followed by a post-centre self-assessment rating. The level of R^2 change was used to measure the addition to predictive variance demonstrated by self-assessment. The effect of OFFER only on PERFORM was not significant ($R^2 = .07717$; $F = 3.679$, $df 1, 44$, NS). The change in R^2 (.03111) obtained by entering self-assessment was not significant giving an overall R^2 of .10828 ($F = 2.611$, $df 2, 43$, NS).

The effect of OFFER only on PROMOTE was not significant ($R^2 = .07784$; $F = 3.852$, $df 1, 45$, NS). The entry of self-assessment produced a nonsignificant R^2 change resulting in an overall R^2 of .08584 ($F = 2.066$; $df 2, 44$, NS). Neither the organisational job offer decision nor self-assessment appeared to have any predictive validity for this group of subjects.

Effect of assessment centre on self-assessment Correlations between a self-rating made prior to the assessment centre and self and assessor ratings made at the conclusion of the assessment centre were computed and are displayed in Figure 9.1. The pre-centre self-rating of job performance was significantly related to the post-centre self-rating of job performance ($r = .57$, $p < .01$) and the post-centre self-rating of performance in the assessment centre ($r = .58$, $p < .01$) but not with OFFER ($r = .13$, NS). The average self-rating of job performance before the assessment centre was 7.45 ($SD = .122$, maximum 10, minimum 5) dropping to 6.92 ($SD = .144$, maximum 9, minimum 4) for self-ratings made after the centre. This difference was significant ($t = 4.40$, $df 84$, $p < .001$). The average self-rating of performance in the assessment centre was 6.39 ($SD = .181$, maximum 8, minimum 3), a significant change from the pre-centre self-ratings ($t = 8.93$, $df 82$, $p < .001$).

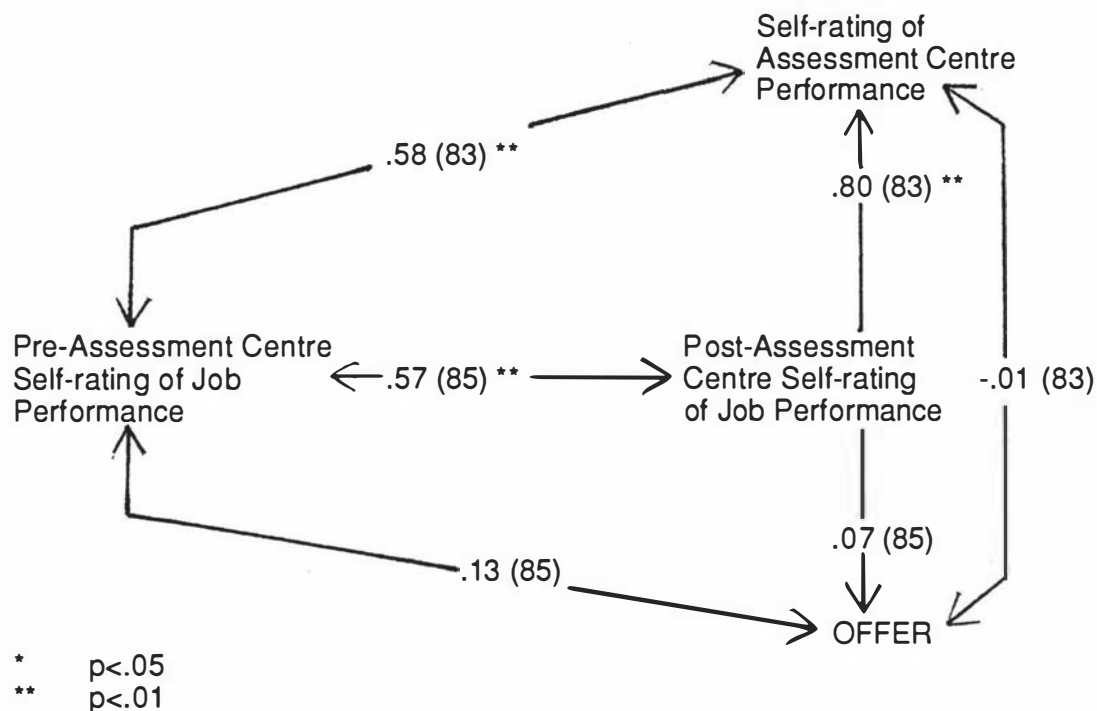


Figure 9.1. Self and Organisational Ratings of Job and Assessment Centre Performance Made Before and After an Assessment Centre. (numbers in brackets show number of cases on which a correlation is based)

Self- and assessor ratings were also made for each of seven work samples used in the assessment centre. All pre- and post self-ratings were significantly and positively related with the exception of the speech summary work sample which had a nonsignificant relationship. No self-rated exercises significantly related to ratings made by assessors therefore the same trends observed in the overall ratings persisted for individual exercises.

The post-centre self-rating of performance in the assessment centre and the self-rating of future job performance were strongly related ($r = .80, p < .001$).

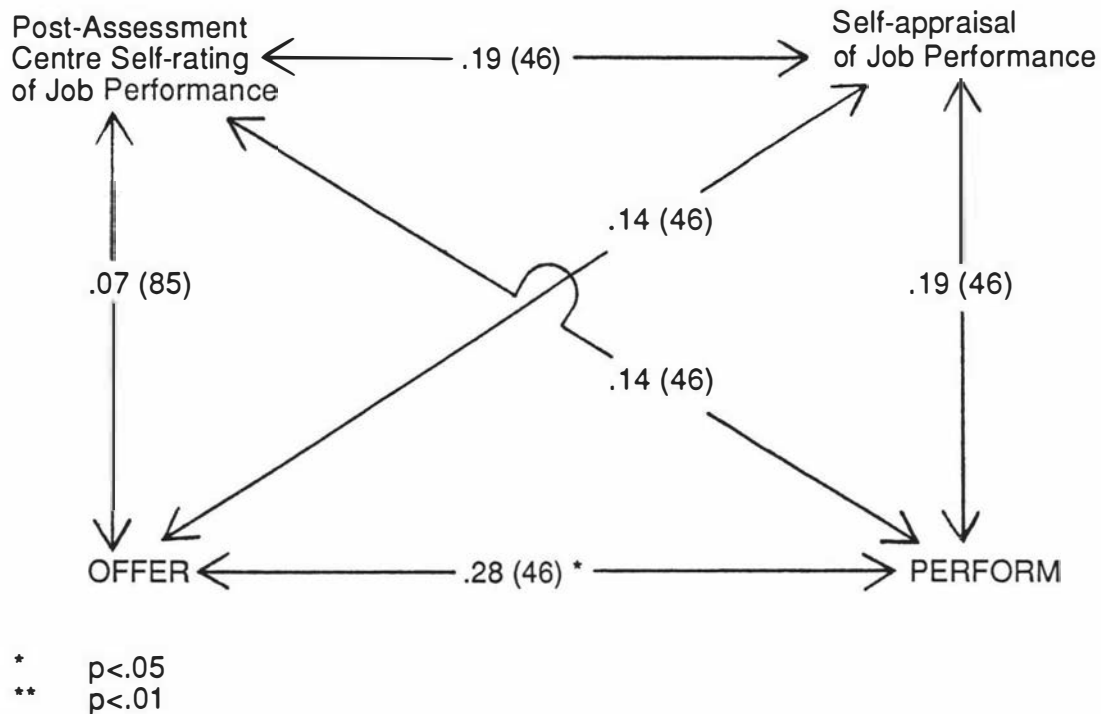


Figure 9.2. Relationships Between Self and Organisational Assessments of Job Performance Made During Selection and When Employed. (numbers in brackets show number of cases on which a correlation is based)

Self-assessment and organisational assessment of job performance. Figure 9.2 shows the relationships between OFFER, a self-rating made during selection, a self-rating made one year later, and PERFORM. The relationships were generally weak with the assessor rating being the only significant predictor of PERFORM ($r = .28$, $p < .05$). The lack of agreement between self and superior ratings of job performance observed during selection persisted when performance was appraised on the job.

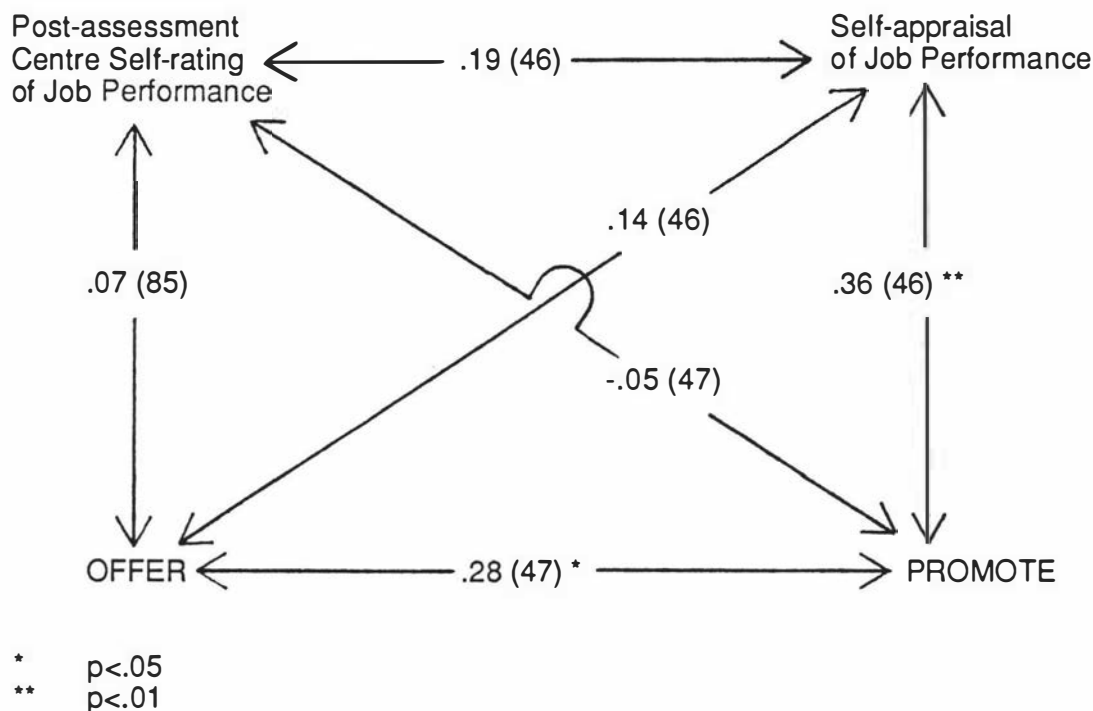


Figure 9.3. Relationships Between Self-Assessments and Organisational Promotion Decisions Made During Selection and When Employed. (numbers in brackets show number of cases on which a correlation is based)

Self-assessment and promotion decisions. When a promotion decision was used as a criterion instead of performance appraisal, the predictive relationship of self-assessment did not change significantly ($r = -.05$, NS). The major difference between Figures 9.2 and 9.3 is the relationship between self-appraised job performance and PROMOTE. Self-appraisal was significantly related to PROMOTE ($r = .36$, $p < .01$) but not to PERFORM (Table 9.2). Although promotion was based on similar information provided by the performance appraisal, self-appraisal appeared to have more in common with the PROMOTE criteria. OFFER was significantly related to PROMOTE ($r = .28$, $p < .05$).

Moderator analysis. The relationship between concurrent and predictive self/organisational ratings of job performance were further examined for the possible

moderating effects of age, gender, intelligence, social desirability, full time work experience, and education. The predictive relationships specifically considered were those between pre- and post-centre self-ratings of job performance, post-centre self-ratings and self-appraised job performance, post-centre self-ratings and PERFORM, and post-centre self-ratings and PROMOTE. The concurrent relationships examined were post-centre self-ratings and OFFER, self-appraised job performance and PERFORM, and self-appraised job performance and PROMOTE.

The large number of relationships (seven) and moderators (six), giving 42 comparisons, suggests some may be significant due to chance. The purpose of the analysis, however, was to examine the moderators across a number of relationships to isolate any consistent moderator effects. No such consistency was found for any of the potential moderators with only one significant between sub-group difference detected (see Appendix 13).

For applicants with no post-graduate qualification ($n = 47$) the correlation between post-centre self-assessment and OFFER was $-.10$, whereas for applicants with (or completing) a post-graduate qualification the correlation was $.34$ ($p < .05$). A post-graduate qualification moderator effect was not found for any other comparisons. Correlations between organisational criteria, self-assessment criteria, and potential moderators were also calculated to identify any main effects. No correlations were significant (two tailed test).

Job acceptance - self-assessment ratings. Self-ratings of assessment centre performance, future job performance, and individual work sample performance were correlated with applicants' acceptance or rejection of job offers. No significant relationships were found suggesting applicant perceptions of performance in the

assessment centre and future job performance had no effect on acceptance or rejection of a job offer.

Discussion

In Study B self-assessment demonstrated no predictive utility as a predictor of organisational work performance criteria. In fact there appears to be no relationship between self-assessments and organisational assessments of work performance, a finding that was generally consistent across predictive and concurrent relationships. Secondly, the assessment centre significantly reduced the level of self-assessed performance in much the same way realistic job previews have been shown to reduce initial job expectations (Premack & Wanous, 1985). Thirdly, there were no significant moderator effects on the relationship between self-assessments and organisational estimates of work performance.

The consistency with which different self-assessments of job performance were related to one another and unrelated to organisational measures suggests a major difference between the two rating sources. Other research has shown that different sources of rating information (self, supervisor, and peer) use different dimensions for assessment (Klimoski & London, 1974; Staley & Shockley-Zalaback, 1986). In addition, Thornton (1980) suggests that incumbents' views of their own job performance are significantly different from the views held by others. Parker, Taylor, Barrett, and Martens (1959) suggested such differences could be due to employees lacking insight into their own strengths and weaknesses or superiors being unwilling or unable to recognise strengths and weaknesses in their subordinates. Rating errors involving personal biases, harsh or lenient judgments, or the tendency to assign average ratings (Landy & Farr, 1983) could also affect the accuracy of such judgments.

The effect of the assessment centre on self-assessment was a significant drop in the overall average self-ratings and five of the seven individual ratings of work sample performance. Schmitt, Ford, and Stults (1986) also showed the effect of taking part in an assessment centre on self-ratings for a large group of participants ($n = 1693$) was small but significant. That particular investigation, however, involved self-ratings of skill dimensions, not overall future job performance or specific exercise performance. It may be that the lowering of average self-ratings is similar to the "vaccination" against unrealistic expectations demonstrated in realistic job previews (Popovich & Wanous, 1982). Participation gives the candidate experiential information so increasing the chances of accurate self-judgments (Bandura, 1977b). For example, not having an understanding of job content or item statements in common with the tester or organisation can negatively affect the association between self-ratings and organisational criteria (Primoff, 1980). Receiving of new information encourages a reassessment of personal ability (Bandura, 1982a), as has been demonstrated in self-selection trends following the acquisition of experiential job information (Downs, Farr, & Colbeck, 1978).

The difference between pre- and post-centre self-assessments in this setting was not reflected in the relationship these self-assessments had with PERFORM and PROMOTE. Pre- and post-centre ratings similarly were poor predictors.

Interestingly, there also was no increase in accuracy between self-appraisals made after the experience of 12 months employment and PERFORM but the relationship between self-appraised on-job performance and PROMOTE was significant (Figure 9.3). Twelve months employment would allow an advisory officer ample opportunity to gain experiential information regarding their work performance. That such knowledge

seems to be reflected in the relationship self-assessment had with the promotional measure, and not the performance ratings, is puzzling. It should be noted that promotional decisions were made by a committee of managers. Performance ratings were made individually by immediate superiors only. It has been suggested that ratings made by a single person can be problematic in that they are subject to individual bias being based on less information than a committee process (Thornton & Byham, 1982).

The large number of relationships between self-assessment and organisational criteria provided the opportunity to evaluate potential moderators in terms of consistency of effect. Overall there was no such consistency with only one significant sub-group difference in a set of 42 comparisons. Following the assessment centre experience, candidates having or completing post-graduate qualifications made self-assessments of future job performance that were positively related to OFFER ($r = .34, p < .05$), and significantly different to similar self-assessments made by candidates without post-graduate experience. Education has been shown to have both a positive effect (e.g., Kooker, 1974) and no effect (e.g., Brief, Aldag, & van Sell, 1977), on self-assessment accuracy. The isolated result found in this study cannot be regarded as giving support to either of the previously reported findings.

The self-selection trend reported by Downs et al. (1978), was not evident in this study. No significant relationship between self-assessment and acceptance or rejection decisions was detected. It is possible that job acceptance decisions were also affected by economic and job market factors outside the selection process so obscuring any such relationship.

An interesting feature of Study B in this setting was the lack of association between self-assessments and organisational assessments of work performance, particularly those measures obtained during the selection procedures. It may be that in this setting both parties (self and organisation) were credible sources of performance information. Each provided independent variance in the prediction of some criterion. Neither gave a complete or totally accurate measure of job performance. If this is so, the best approach to selection in similar circumstances may be a collaborative one. Both parties contributing to the process could share their independently generated assessments and together attempt to reach an optimal employment decision.

CHAPTER TEN

SELF-ASSESSMENT IN NEW ZEALAND PERSONNEL SELECTION

This research project set out to evaluate self-assessment as a personnel selection predictor in New Zealand institutional personnel selection. In specific terms this involved the following:

1. An evaluation of the acceptability of self-assessment as a selection tool to New Zealand practitioners.
2. An evaluation of the unique predictive variance offered by self-assessment in applied settings above that provided by selection methods used in New Zealand organisations.
3. An evaluation of possible moderator effects on the self-assessment - work criteria relationships. Potential moderators included cognitive ability, gender, educational level, age, social desirability, locus of control, level of job information, self-assessment experience, and whether or not self-assessment requests are made in "compared to others" or absolute terms.

In short, the results showed that few New Zealand practitioners accepted self-assessment as a useful selection method, that the validity offered by self-assessment was poor, and that no consistent moderator effects amongst those tested were identified. These results are discussed in terms of their implications for New Zealand personnel selection.

The Acceptability of Self-Assessment in New Zealand

For a selection technique to be of any use it has to be acceptable to those who will use it. The few willing participant organisations and the high level of attrition amongst established research settings showed that self-assessment as a selection predictor had little attraction for New Zealand practitioners. The real reason for organisations refusing to participate is difficult to establish, however. Only two organisations specifically stated they had little faith in self-assessment as a selection method. Companies may have refused access as part of a policy applying to any research, or alternatively, because they believed self-assessment was of little use in personnel selection. It is also possible that some practitioners felt their current procedures were adequate and it was not worth considering any alternative selection procedure.

The withdrawal of settings in which self-assessment was implemented occurred for a variety of stated reasons. Perhaps most important in terms of the acceptability issue were the two organisations which withdrew from the project because of negative reactions from candidates. The practitioners in these companies were concerned that such a response would prejudice their recruitment programmes in areas where expertise was in short supply. It has generally been assumed that asking those normally judged by others to participate in the assessment process is a positive move. Such an approach was felt to encourage an atmosphere of trust rather than confrontation (e.g., Bassett & Meyer, 1968). The reactions from some applicants suggested that using self-assessment in a selection setting may not necessarily have been appreciated by all candidates. Candidates in one instance remarked that assessment was the personnel officer's job, not theirs. A self-assessment request was regarded as beyond the bounds of what was expected in New Zealand selection procedures. This response may be linked to New Zealand workers avoiding

participation in decision making or sharing responsibility (Hines, 1974). Asking individuals to predict their own performance involves the self-assessor taking responsibility for his or her response, a responsibility some applicants may prefer to avoid.

An unwillingness to participate in decision making may also be linked to certain cultural values. During the course of the research, personnel officers in two government departments described to the researcher how some Maori tribal groups have participated in selection procedures. Instead of different applicants directly applying for positions, tribal elders approached the prospective employers, evaluated the job themselves, and then recommended a tribal member for the position. In effect, prospective employees were selected by their tribal group because the employer was under some obligation to accept the decision of the elders. Such a procedure could be described as a special case of peer selection and illustrates a rather unique approach to selection decision making.

Another important reason given for withdrawing from the project was the reluctance of a newly appointed manager to continue the research initiated by a predecessor. This was the most common reason for withdrawal, a trend that emphasises the crucial role of the contact person who takes responsibility for research being conducted and completed in the organisation. Whether this tendency is peculiar to New Zealand is impossible to say, especially when the experience of such practical events is hardly ever reported in the research literature (Kulka, 1982). Access for researchers to real life settings is difficult enough. It seems that, at least in New Zealand organisations, the permanence of access rarely extends beyond the tenure of the contact person.

Of the four organisations which remained in the project for its full term, the Polytechnic is the only one which continued to use self-assessment. A version of this organisation's self-completed reference form was already part of the selection process prior to the investigation. The other three organisations did not use self-assessment prior to the project and did not use it afterwards. The organisational decision to stop using self-assessment was made before these three organisations received any feedback on the validity of the approach. Self-assessment in these settings was not acceptable as a selection predictor.

Overall, it seems that self-assessment was not attractive enough to New Zealand practitioners to gain organisational access to allow a complete assessment of acceptability. In those settings where self-assessment was implemented it did not survive personnel changes or reorganisation and did not continue beyond the life of the research project.

The Validity of Self-Assessment in New Zealand

The validity demonstrated by self-assessment as a predictor was poor. When summarised as population correlations (Hunter, Schmidt, & Jackson, 1982) calculated across studies for each of the main criteria, no validity coefficient was significant. For supervisory ratings of work performance the population correlation was .05 ($n = 4$, $s^2 = .0078$), for promotion $-.14$ ($n = 2$, $s^2 = .0149$), for duration of absence $.00$ ($n = 2$, $s^2 = .1101$). The zero order correlation for the absence frequency criterion was $.11$. In the seasonal worker studies only two turnover groups were involved. The population correlation for these two samples was $.09$ ($s^2 = .0279$). Correlational

2

S^2 represents the standard deviation of the calculation which in the case of the population correlation is the frequency of weighted average squared error (Hunter, Schmidt & Jackson, 1982).

descriptions are not possible for self-assessment - turnover relationships where the criterion involved three nominal categories.

There are a number of possible reasons why the validity of self-assessment demonstrated in this investigation was poor. It would seem that one of the major factors affecting self-assessment validity was the selection situation itself. Generally people apply for jobs because they want them, an aspiration which may conflict with a request for accurate self-assessment information. Accurate self-appraisal may be a disadvantage to the candidate if it shows that they are unsuitable for a job.

Often the reaction of individuals toward others who control desired resources is to engage in various self-presentation strategies (Arkin, 1980). For instance, applicants may seek to create an impression of competence or even intentionally deceive company personnel with the aim of becoming liked or seen as particularly worthy of a job. As Ash (1980) and Elliot (1981) suggest, the requirements of a personnel selection process encourage socially desirable and sometimes distorted responses. Inflated or lenient self-assessments were apparent in all samples involved in the present study. Few applicants rated themselves below average and self-assessment distributions were generally skewed to the highly positive end of the scales used.

The propensity for applicants to inflate their self-ratings could have resulted in what Wicklund and Gollwitzer (1983) termed "antivalidity" in three of the project studies. In these cases the relationship between self-assessment and performance criteria was negative. Those giving themselves high self-ratings at selection were more likely to receive low performance appraisals or be promoted more slowly than their peers. The theory of symbolic self-completion, suggests that the tendency to be positive should be expected primarily from those who are least competent and committed to a certain

self-description, or in this case, a certain job (Wicklund & Gollwitzer, 1983). Positive self-descriptions were related to an absence of expertise. In other words, those who want the job and feel they are not well qualified are most likely to provide inflated self-assessments. Instances of this trend have been observed in other settings (e.g., Primoff, 1980).

The poor predictive validity of self-assessment may also be partly due to what is being predicted. The supervisor's rating as the actual criterion is possibly accounting for only part of the variance required in an ultimate criterion of work performance. Variance supplied by self-assessment may be correcting for some of the deficiencies of the actual criterion by supplying additional independent information. The low relationships reported between self- and supervisory assessments of performance can be interpreted as a demonstration of independence, not poor validity (e.g., Brief, et al., 1977; McEnery & McEnery, 1987). An example of self and others ratings being reasonably independent was illustrated in the government department study. Of all ratings collected during the selection procedure no pre- or post assessment centre self-ratings were related to the organisational job offer but all self-ratings were strongly related to each other. This pattern did not persist when predictive comparisons were made, a trend that may have more to do with the nature of the criteria than a lack of independence.

It is worth noting that the presence of various measurement conditions did not appear to strengthen the association between self-assessment and work performance criteria. According to Mabe and West (1982), measurement conditions exert considerable influence on the validity of self-assessment. In the studies analysed using organisational performance appraisals as criteria, self-assessment requests were matched, required evaluations of performance not ability, were in comparative terms,

and were made relative to a target group. The population correlation reported above for these studies was .05 ($n = 4$, $s^2 = .0078$), a substantial difference from .63, the validity Mabe and West expected when any four measurement conditions are present. In addition, self-assessment experience appeared to have little moderating effect on the validity of self-assessment across all studies analysed.

The evidence describing the validity of self-assessment in selection indicates that the use of this technique may be limited to those situations where it is more important for the assessor to be accurate than to use the assessment to achieve an end (which may conflict with accuracy). The success demonstrated in education settings, vocational guidance, and other applications has not been reflected in New Zealand personnel selection.

Moderator Effects on the Validity of Self-Assessment

Amongst the potential moderators investigated, few demonstrated significant effects. Those effects that were significant may in isolation appear interesting, but in the context of a large number of comparisons only one of the moderators evaluated produced an effect in more than one study.

Educational level was identified as a moderator in a seasonal worker sample, a student nurse sample, and in a concurrent relationship in the government officer sample (Study B only in each case). In general terms the effect was in favour of cases who had obtained relatively better educational qualifications than their peers. Age was identified as a moderator only in the first seasonal worker study. The self-assessments made by older workers were positively related to criteria measuring performance, absenteeism, and turnover. These effects did not persist in the second

study conducted in the same setting nor were they found in any of the other studies in the project.

Two moderators which were expected to affect the validity of self-assessment did not. Firstly, the poor predictive validity demonstrated by self-assessment was not related to the amount of job knowledge possessed by applicants as suggested by Bandura (1982a). Across the different studies, selection procedures and reported individual job knowledge varied widely yet there was no noticeable trend toward applicants with more extensive information being better at self-assessing job performance. Secondly, the expectation that self-assessments requested in relative terms were more accurate than requests which did not ask for a comparative assessment (Mabe & West, 1982) was also not supported. In the seasonal worker sample comparative and absolute requests were equally poor.

Overall, there was little consistency in moderator effects identified across relationships within samples and across different studies. None of the potential moderators investigated were present across even 10% of the comparisons made. There are a number of possible reasons why this was so. Firstly, moderator effects may be minimal for self-assessment - work performance criteria relationships. Brief et al., (1977) reported findings supporting this possibility and reviews have found the direction and presence of such effects inconsistent (Mabe & West, 1982). Secondly, in those studies where sample size allowed an investigation of moderator effects, the sub-group sizes involved in some cases were very small so reducing statistical power. Thirdly, the distribution of the self-assessment ratings in all studies was skewed to the upper end of the rating scale. This may have been due to the nature of the scale such as in the case of nursing trainees. When asked whether they would complete the course, all applicants answered yes. Generally, applicants rarely rated themselves

as below average and as a result the variability of self-assessment ratings was reduced, further lessening the chance of detecting significant moderator effects.

The possibility that an applicant's preference for a particular job in the seasonal worker sample was affecting their self-assessment was also investigated. It was not expected that job preference would act as a moderator but would be positively related to self-assessment. More preferred jobs would receive higher self-assessments. A significant level of common variance between job preference and self-assessment was detected ($r = .18, p < .001$). Where some choice existed, such as in the seasonal worker setting, applicants appear to have biased their self-judgments towards more preferred jobs. The effect of job preference was restricted to the self-assessment ratings. Job preference had no significant variance above that of self-assessment to offer in the prediction of performance, turnover, and absence criteria in that particular setting.

Possibilities for Further Research and Practice in New Zealand

The results and experience of conducting the present investigation suggests a number of possibilities for further research. Firstly, the likelihood that the poor validity of self-assessments in selection situations is due to committed applicants with self-perceived skill weaknesses hiding deficiencies with high and inaccurate ratings, is worthy of investigation. Particular attention needs to be paid to developing procedures to reduce or avoid such antivalidity effects. Self-completion theory suggests encouraging applicants to abandon self-symbolising for self-reflection would increase the accuracy of their self-reports. Applicants could be encouraged to evaluate both their strengths and weaknesses so clarifying an overall self-assessment. Two settings in the present investigation felt self-assessment was exceptionally useful for identifying a candidate's training needs. It may also be valuable to identify those

individuals or situations that prevent such self-reflection. For example, there is a chance that when the assessment is for the benefit of the self-assessor, dissimulation does not occur. In addition, antivalidity or leniency effects may be absent in those situations where the job seeker has a degree of choice between positions in different organisations.

Secondly, approaches to incorporating expectations of validation in the minds of applicants in real world selection settings could also be examined. Previous work had examined the effect of presence and absence of this condition in laboratory settings (Farh & Werbel, 1986), but little further work has been done. Different modes of presentation or even training procedures involving opportunities to reward accurate self-evaluation could be tried. As with any investigation aimed at practical application, the ethical and functional aspects of any proposed procedure must be part of the inquiry.

Thirdly, while self-assessment may be of little value to an employer as a predictor, the effect of an applicant's self-evaluative processes cannot be ignored. Consider, for instance, the drastic reduction in productivity gains associated with the use of valid selection techniques when highly successful applicants refuse job offers (Murphy, 1986). Applicants are not simply swayed by recruiting practices but seem to make acceptance decisions based on job attributes (Powell, 1984), perceptions of their own capacity to perform various tasks (Wheeler, 1984), and possibly job preference.

It has been demonstrated in this investigation that self-assessment as a predictor has little to offer in the context of current selection procedures. It is possible that a more participative style of personnel selection would be feasible in certain jobs giving an employer and a candidate an opportunity to collaborate in a joint decision. As

suggested by Robertson and Makin (1986), this would involve creating a selection process that maximises the amount of information shared by the employer and candidate. These authors state, "it seems curious that the person with most to lose from a poor decision is not given a more central role in decision making" (p. 55).

Such a process would also depend on an employer's acceptance of a bilateral approach to personnel selection. New Zealand practitioners, probably in common with those in other countries, appear to see the decision making in personnel selection as their exclusive responsibility. The role of the candidate is to provide information and react to an employer's offer. Increasing unemployment in many work areas supports such unilateral decision making. Employers are the purchasers of applicant skills in a job market favouring the buyer. Even in those areas where there is a shortage of individuals the emphasis appears to be on ensuring the employer still has a choice of candidates. Employers retain almost complete control of the selection process.

For selection decisions to become truly collaborative employers will have to relinquish some of their decision making power and focus more on processes which use the resources of both parties in the selection process. It is possible that as a result many of the demand characteristics inherent in the traditional unilateral approach would be substantially reduced. It is almost as if those involved in personnel selection have to undergo a similar process to that seen in the early days of work motivation theories. The so called theory of "scientific" management preached workers work only for money. The Hawthorne studies and other theoretical efforts revealed that workers cannot only think for themselves but will actively participate to improve productivity. Those involved in selection research and practice also have to realise that applicants can and do think for themselves if the conditions under which selection takes place will allow them to do so.

REFERENCE LIST

- Aberth, J. (1986). Pre-employment testing is losing favour. Personnel Journal, 65, 96-104.
- Abrahams, N. M., & Alf, E. (1972). Pratfalls in moderator research. Journal of Applied Psychology, 56, 245-251.
- Adorno, T. W., Frenkel-Brunswick, E., Levinson, D. J., & Sanford, R. N. (1950). The Authoritarian Personality. New York: Harper.
- Alderfer, C. P., & McCord, C. G. (1970). Personal and situational factors in the recruitment interview. Journal of Applied Psychology, 54, 377-385.
- Andersen, S. M. (1984). Self-knowledge and social inference: II. The diagnosticity of cognitive/affective and behavioural data. Journal of Personality and Social Theory, 46, 294-307.
- Andersen, S. M., & Ross, L. (1984). Self-knowledge and social inference: I. The impact of cognitive/affective and behavioural data. Journal of Personality and Social Theory, 46, 280-293.
- Anderson, C. D., Warner, J. L., & Spencer, C. C. (1984). Inflation biases in self-assessment examinations: Implications for valid employee selection. Journal of Applied Psychology, 69, 574-580.
- Anderson, J. R. (1982). Acquisition of cognitive skill. Psychological Review, 89, 369-406.
- Anderson, L. R. (1983). Management of the mixed cultural work-group. Organisational Behaviour and Human Performance, 31, 303-330.
- Anderson, N., & Shackleton, V. (1986). Recruitment and selection: A review of developments in the 1980's. Personnel Review, 15, 19-26.
- Anstey, E. (1971). The civil service administrative class: A follow-up of post-war entrants. Occupational Psychology, 45, 27-43.

- Anstey, E. (1977). A 30-year follow-up of the CSSB procedure, with lessons for the future. Journal of Occupational Psychology, 50, 149-159.
- Arkin, R. M. (1980). Self-presentation. In D. M. Wegner & R. R. Vallacher (Eds.), The Self in Social Psychology. New York: Oxford University Press.
- Arnold, H. J. (1982). Moderator variables: A clarification of conceptual, analytic and psychometric issues. Organisational Behaviour and Human Performance, 29, 143-174.
- Arnold, H. J., & Evans, M. G. (1979). Testing multiplicative models does not require multiplicative scales. Organisational Behaviour and Human Performance, 24, 41-59.
- Arnold, H. J., & Feldman, D. C. (1981). Social desirability response bias in self-report choice situations. Academy of Management Journal, 24, 377-385.
- Arnold, L., Willoughby, T. E., & Calkins, E. V. (1985). Self-evaluation in undergraduate medical education: A longitudinal perspective. Journal of Medical Education, 60, 21-28.
- Arvey, R. D., & Campion, J. E. (1982). The employment interview: A summary and review of previous research. Personnel Psychology, 35, 281-322.
- Arvey, R. D., Miller, H. E., Gould, R., & Burch, P. (1987). Interview validity for selecting sales clerks. Personnel Psychology, 40, 1-12.
- Ash, A. A. (1980). Self-assessment of five types of typing ability. Personnel Psychology, 33, 273-282.
- Ash, A. A., & Levine, E. L. (1985). Job applicant training and work experience evaluation: An empirical comparison of four methods. Journal of Applied Psychology, 70, 572-576.
- Ash, R. A., Levine, E. L., & Edgell, S. L. (1979). Exploratory study of a matching approach to personnel selection: The impact of ethnicity. Journal of Applied Psychology, 64, 35-41.
- Asher, J. J. (1972). The biographical item: Can it be improved? Personnel Psychology, 25, 251-269.

- Asher, J. J., & Sciarrino, J. A. (1974). Realistic work sample tests: A review. Personnel Psychology, 27, 519-533.
- Bailey, R. C., & Bailey, K. G. (1974). Self-perceptions of scholastic ability at four grade levels. The Journal of Genetic Psychology, 124, 197-212.
- Baird, L. S. (1977). Self and superior ratings of performance: As related to self-esteem and satisfaction with supervision. Academy of Management Journal,
- Bandura, A. (1977a). Social Learning Theory. Englewood Cliffs, New Jersey: Prentice-Hall.
- Bandura, A. (1977b). Self-efficacy: Towards a unifying theory of behavioural change. Psychological Review, 84, 191-215.
- Bandura, A. (1978). The self system in reciprocal determinism. American Psychologist, 33, 344-358.
- Bandura, A. (1982a). Self-efficacy mechanism in human agency. American Psychologist, 37, 122-147.
- Bandura, A. (1982b). The self and mechanisms of agency. In J. Suls (Ed.), Psychological Perspectives on the Self (Vol. 1, pp. 3-39). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Bandura, A., Adams, N. E., & Beyer, J. (1977). Cognitive processes mediating behavioural change. Journal of Personality and Social Psychology, 35, 125-139.
- Bandura, A., & Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. Journal of Personality and Social Psychology, 45, 1017-1028.
- Barnes, C. L., & Jamieson, B. D. (1977). Reactions to job characteristics among Samoan, Pakeha, and Maori employees. New Zealand Psychologist, 6, 14-24.

- Barrett, G. V., Phillips, J. S., & Alexander, R. A. (1981). Concurrent and predictive validity designs: A critical reanalysis. Journal of Applied Psychology, 66, 1-6.
- Bass, B. M., & Burger, P. (1979). Assessment of Managers: An International Comparison. New York: Free Press.
- Bassett, G. A., & Meyer, H. H. (1968). Performance appraisal based on self-review. Personnel Psychology, 21, 421-430.
- Bauman, K. E., & Dent, C. W. (1982). Influence of an objective measure on self-reports of behaviour. Journal of Applied Psychology, 67, 623-628.
- Baumeister, R., & Jones, E. E. (1978). When self-presentation is constrained by the target's knowledge: Consistency and compensation. Journal of Personality and Social Psychology, 36, 608-618.
- Baxter, J. C., Brock, B., Hill, P. C., & Rozelle, R. M. (1981). Letters of recommendation: A question of value. Journal of Applied Psychology, 66, 296-301.
- Belt, J. A., & Holden, P. B. (1978). Polygraph usage among major U.S. corporations. Personnel Journal, 57, 80-86.
- Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology (Vol. 6, pp. 1-62). New York: Academic Press.
- Ben-Shakhar, G., Bar-Hillel, M., Bilu, Y, Ben-Abba, E., & Flug, A. (1986). Can graphology predict occupational success? Two empirical studies and some methodological ruminations. Journal of Applied Psychology, 71, 645-653.
- Berkowitz, L., & Daniels, L. R. (1963). Responsibility and dependency. Journal of Abnormal and Social Psychology, 66, 664-669.
- Betz, N. E., & Hackett, G. (1981). The relationship of career-related self-efficacy expectations to perceived career options in college women and men. Journal of Counseling Psychology, 28, 399-410.

- Bills, R. E., Vance, E. L., & McLean, O. S. (1951). An index of adjustment and values. Journal of Consulting Psychology, 15, 257-261.
- Boehm, V. R. (1982). Are we validating more but publishing less? The impact of governmental regulation on published validation research: An exploratory investigation. Personnel Psychology, 35, 175-187.
- Borman, W. C. (1978). Exploring upper limits of reliability and validity in job performance ratings. Journal of Applied Psychology, 63, 135-144.
- Bower, G. H., & Gilligan, S. G. (1979). Remembering information related to one's self. Journal of Research in Personality, 13, 420-432.
- Bray, D. W., & Grant D. L. (1966). The assessment centre in the measurement of potential for business management. Psychological Monographs: General and Applied, 80 (Whole No. 625).
- Breaugh, J. A. (1981). Predicting absenteeism from prior absenteeism and work attitudes. Journal of Applied Psychology, 66, 555-560.
- Brewster-Smith, M. (1966). Explorations in competence: A study of Peace Corps teachers in Ghana. American Psychologist, 21, 555-566.
- Brief, A. P., & Aldag, R. J. (1981). The "self" in work organisations: A conceptual review. Academy of Management Review, 6, 75-88.
- Brief, A. P., Aldag, R. J., & van Sell, M. (1977). Moderators of the relationship between self and superior evaluations of job performance. Journal of Occupational Psychology, 50, 129-134.
- British Psychological Society. (1978). Report of the Working Party on the Sex Discrimination Act. Leicester: Author.
- Brogden, H. E. (1949). When testing pays off. Personnel Psychology, 2, 171-184.

- Brown, I., & Inouye, D. K. (1978). Learned helplessness through modeling: The role of perceived similarity in competence. Journal of Personality and Social Psychology, 36, 900-908.
- Browning, R. C. (1968). Validity of reference ratings from previous employers. Personnel Psychology, 21, 389-393.
- Buckle, P. W. (1985). Self-reported anthropometry. Ergonomics, 28, 1575-1577.
- Bull, P. E. (1974). Should the 16PF be used in personnel selection? New Zealand Psychologist, 3, 11-15.
- Bull, P. E., & Spicer, J. (1983). Actual and potential use of psychologists in New Zealand business. Bulletin Of the New Zealand Psychological Society, 37, 4-5.
- Buros, O. K. (1965). The Sixth Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press.
- Buros, O. K. (1972). The Seventh Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press.
- Bycio, P., Alvares, K. M., & Hahn, J. (1987). Situational specificity in assessment centre ratings: A confirmatory factor analysis. Journal of Applied Psychology, 72, 463-474.
- Byham, W. C., & Wettengel, C. (1974). Assessment centres for supervisors and managers: An introduction and overview. Public Personnel Management, 3, 352-364.
- Byrne, B. M. (1984). The general/academic self-concept nomological network: A review of construct validation research. Review of Educational Research, 54, 427-456.
- Byrne, B. M., & Shavelson, R. J. (1986). On the structure of the adolescent self-concept. Journal of Educational Psychology, 78, 474-481.
- Callender, J. C., & Osburn, H. G. (1982). Another view of progress in validity generalisation: Reply to Schmidt, Hunter, and Pearlman. Journal of Applied Psychology, 67, 846-852.

- Campbell, J. P., Dunnette, M. D., Lawler, E. E., & Weick, K. E. (1970). Managerial Behaviour, Performance, and Effectiveness. New York: McGraw-Hill.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, *56*, 81-105.
- Campion, M. A. (1983). Personnel selection for physically demanding jobs: Review and recommendations. Personnel Psychology, *36*, 527-550.
- Carroll, S. J., & Nash, A. N. (1972). Effectiveness of a forced-choice reference check. Personnel Administration, *35*, 42-46.
- Cascio, W. F. (1987). Costing Human Resources: The Financial Impact of Behaviour in Organisations (2nd ed.). Boston: Kent.
- Cascio, W. F., & Silbey, V. (1979). Utility of the assessment centre as a selection device. Journal of Applied Psychology, *64*, 107-118.
- Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. (1980). Handbook for the Sixteen Personality Factor Questionnaire (16PF). Champaign, Illinois: IPAT.
- Ceci, S. J., & Peters, D. (1984). Letters of reference: A naturalistic study of the effects of confidentiality. American Psychologist, *39*, 29-31.
- Chadwick-Jones, J. K., Nicholson, N., & Brown, C. (1982). Social Psychology of Absenteeism. New York: Praeger Publishers.
- Chiu, L. (1975). Influence of student teaching on perceived teaching competence. Perceptual and Motor Skills, *40*, 872-874.
- Christensen, L. (1974). The influence of trait, sex, and information on accuracy of personality assessment. Journal of Personality Assessment, *38*, 130-135.
- Cleary, T. A., Humphreys, L. G., Kendrick, S. A., & Wesman, A. (1975). Educational uses of tests with disadvantaged students. American Psychologist, *30*, 15-41.

- Cleff, S. H., & Hecht, R. M. (1971). Job/man matching in the 70's. Datamation, 17, 22-27.
- Cohen, J., & Cohen, P. (1983). Applied Multiple Regression/Correlation Analysis for the Behavioural Sciences (2nd ed.). Hillsdale, New Jersey: Lawrence Erlbaum.
- Colarelli, S. M. (1984). Methods of communicating and mediating processes in realistic job previews. Journal of Applied Psychology, 69, 633-642.
- Cronbach, L. J., & Gleser, G. C. (1965). Psychological Tests and Personnel Measurement. Urbana, Illinois: University of Illinois Press.
- Crowne, D. P., & Marlowe, D. (1964). The Approval Motive. New York: Wiley.
- Cunningham, R. B., & Olshfski, D. F. (1985). Evaluating task leadership: A problem for assessment centres. Public Personnel Management, 14, 293-299.
- D'Augelli, A. R. (1973). The assessment of interpersonal skills: A comparison of observer, peer, and self-ratings. Journal of Community Psychology, 1, 177-179.
- Dakin, S., & Armstrong, J. S. (1986). Forecasting job performance: A comparison of expert opinion and research findings. Manuscript submitted for publication.
- Dakin, S., & Inkson, K. (1985, May). Apprentice selection: Time for a change of approach. Management, pp. 56-65.
- Dalton, D. R., Krackhardt, D. M., & Porter, L. W. (1981). Functional turnover: An empirical assessment. Journal of Applied Psychology, 66, 716-721.
- Dalton, D. R., Todor, W. D., & Krackhardt, D. M. (1982). Turnover overstated: The functional taxonomy. Academy of Management Review, 7, 117-123.
- Davids, A., & Pildner, H. (1958). Comparison of direct and projective methods of personality assessment under conditions of motivation. Psychological Monographs, 11 (Whole No. 11).

- Davis, K. R. (1984). A longitudinal analysis of biographical subgroups using Owen's developmental integrative model. Personnel Psychology, 37, 1-14.
- de Wolff, C. J., & van den Bosch, G. (1984). Personnel selection. In P. J. D. Drenth, H. Thierry, P. J. Willems & C. J. de Wolff (Eds.), Handbook of Work and Organisational Psychology (Vol. 1, pp. 289-311). Chichester: John Wiley.
- Dean, R. A., & Wanous, J. P. (1984). Effects of realistic job previews on hiring bank tellers. Journal of Applied Psychology, 69, 61-68.
- Deaux, K. (1979). Self-evaluations of male and female managers. Sex Roles, 5, 571-580.
- DeNisi, A. S., & Shaw, J. B. (1977). Investigation of the uses of self-reports of abilities. Journal of Applied Psychology, 62, 641-644.
- Department of Statistics (1981). New Zealand Census of Population and Dwellings: Volume 4, Labour Force. Wellington, New Zealand: Author.
- Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. Behavioural Science, 19, 1-15.
- Dicken, C. (1969). Predicting the success of Peace Corps community development workers. Journal of Consulting and Clinical Psychology, 33, 597-606.
- Dougherty, T. W., Ebert, R. J., & Callender, J. C. (1986). Policy capturing in the employment interview. Journal of Applied Psychology, 71, 9-15.
- Downs, S. (1984). Testing Trainability. Manuscript submitted for publication.
- Downs, S., Farr, R. M., & Colbeck, L. (1978). Self-appraisal: A convergence of selection and guidance. Journal of Occupational Psychology, 51, 271-278.
- Drenth, P. J. D. (1984). Personnel appraisal. In P. J. D. Drenth, H. Thierry, P. J. Willems & C. J. de Wolff (Eds.), Handbook of Work and Organisational Psychology (Vol. 1, pp. 573-596). Chichester: John Wiley.

- Dunnette, M. D. (1963a). A note on the criterion. Journal of Applied Psychology, 5, 251-254.
- Dunnette, M. D. (1963b). A modified model for test validation and selection research. Journal of Applied Psychology, 5, 317-323.
- Dunnette, M. D. (1972). Comments on Abrahams and Alf's "pratfalls in moderator research". Journal of Applied Psychology, 56, 252-256.
- Dunnette, M. D., & Borman, W. C. (1979). Personnel selection and classification systems. Annual Review of Psychology, 30, 477-525.
- Dunnette, M. D., & Maetzold, J. (1955). Use of a weighted application blank in hiring seasonal employees. The Journal of Applied Psychology, 39, 308-310.
- Eberhardt, B. J., & Muchinsky, P. M. (1982a). An empirical investigation of the factor stability of Owen's Biographical Questionnaire. Journal of Applied Psychology, 67, 138-145.
- Eberhardt, B. J., & Muchinsky, P. M. (1982b). Biodata determinants of vocational psychology: An integration of two paradigms. Journal of Applied Psychology, 67, 714-727.
- Edwards, A. L. (1957). The Social Desirability Variable in Personality Assessment and Research. New York: Dryden.
- Edwards, A. L. (1984). An Introduction to Linear Regression and Correlation (2nd ed.). New York: W. H. Freeman.
- Ekpo-Ufot, A. (1976). Self-perceived abilities relevant in the task (SPART): A potential predictor of labour turnover in an industrial work setting. Personnel Psychology, 29, 405-416.
- Ekpo-Ufot, A. (1979). Self-perceived task relevant abilities, rated job performance, and complaining behaviour of junior employees in a government ministry. Journal of Applied Psychology, 64, 429-434.

- Elliot, A. G. P. (1981). Some implications of lie scale scores in real-life selection. Journal of Occupational Psychology, 54, 9-16.
- England, G. W. (1971). Development and Use of Weighted Application Blanks (rev ed.). Minneapolis, Industrial Relations Centre University of Minnesota.
- Evans, M. G. (1985). A monte carlo study of the effects of correlated method variance in moderated multiple regression analysis. Organisational Behaviour and Human Decision Processes, 36, 305-323.
- Eysenck, H. J., & Eysenck, S. B. G. (1970). The Eysenck Personality Inventory. San Diego: Industrial Testing Service.
- Ezekiel, R. S. (1968). The personal future and Peace Corps competence. Journal of Personality and Social Psychology Monograph Supplement, 8(2, Pt 2), 1-26.
- Farh, J., & Werbel, J. D. (1986). Effects of purpose of the appraisal and expectation of validation on self-appraisal leniency. Journal of Applied Psychology, 71, 527-529.
- Farley, J. A., & Mayfield, E. C. (1976). Peer nominations without peers? Journal of Applied Psychology, 61, 109-111.
- Farr, J. L., O'Leary, B. S., & Bartlett, C. J. (1973). Effect of work sample test upon self-selection and turnover of job applicants. Journal of Occupational Psychology, 58, 283-285.
- Farris, G. F. (1971). A predictive study of turnover. Personnel Psychology, 24, 311-328.
- Feldman, J. M., Ah Sam, I., McDonald, W. F., & Bechtel, G. C. (1980). Work outcome preference and evaluation: A study of three ethnic groups. Journal of Cross-Cultural Psychology, 11, 444-468.
- Festinger, L. A. (1954). A theory of social comparison processes. Human Relations, 7, 117-140.

- Finney, J. W., Mitchell, R. E., Cronkite, R. C., & Moos, R. H. (1984). Methodological issues in estimating main and interactive effects: Examples from coping, social support, and stress field. Journal of Health and Social Behaviour, 25, 85-98.
- Flanagan, J. C. (1959). Technical Report (Flanagan Aptitude Classification Tests). Chicago: Science Research Associates, Inc.
- Fluckiger, F. A., Tripp, C. A., & Weinberg, G. H. (1961). A review of experimental research in graphology, 1933-1960 [Monograph]. Perceptual and Motor Skills, 12, 67-90.
- Folger, R., & Belew, J. (1985). Nonreactive measurement: A focus for research on absenteeism and occupational stress. In B. Staw and L. Cummings (Eds.), Research in Organisational Behaviour (Vol. 7, pp. 129-170). Greenwich, Connecticut: JAI Press.
- Foot, H., & Smith M. C. (1986). Student evaluations of academic performance. New Zealand Journal of Educational Studies, 21, 195-202.
- Fourth Estate (1984). The New Zealand Business Who's Who (24th ed.). Wellington, New Zealand: Author.
- Francis, N. (1984, March). Applied psychology in New Zealand industry and commerce. New Zealand Psychological Society Occupational Division Newsletter, pp. 2-3.
- rank, F. D., & Struth, M. R. (1984, March). The self-assessment centre. Training, pp. 56-59.
- Frederick, C. J. (1965). Some phenomena affecting handwriting analysis. Perceptual and Motor Skills, 20, 211-218.
- Fredriksen, N., & Melville, S. D. (1954). Differential predictability in the use of test scores. Educational and Psychological Measurement, 14, 647-656.
- Freeburg, N. E. (1967). The biographical information blank as a predictor of student achievement: A review. Psychological Reports, 20, 911-925.
- Freud, S. (1925). Collected Papers. London: Hogarth.

- Gadzella, B. M., Cochran, S. W., Parham, L., & Fournet, G. P. (1976). Accuracy and differences among students in their predictions of semester achievement. Journal of Educational Research, 70, 75-81.
- Gaugler, B. B., Rosenthal D. B., Thornton, G. C., & Bentson, C. (1987). Meta-analysis of assessment centre validity [Monograph]. Journal of Applied Psychology, 72, 493-511.
- George, D. I. (1984). Self-assessment: An evaluation of potential contribution to personnel selection. Palmerston North, New Zealand: Massey University, Psychology Department.
- George, D. I., & Hyde, K. F. (1984). Graduate Recruitment. Palmerston North, New Zealand: Massey University, Psychology Department.
- George, D. I., & Smith, M. C. (1986a). The Trade and Industry Assessment Centre Experience. Palmerston North, New Zealand: Occupational Psychology Research Associates Ltd.
- George, D. I., & Smith, M. C. (1986b, December). Personnel selection in New Zealand: Does it measure up? Labour and Employment Gazette, pp. 40-42.
- George, D. I., & Smith, M. C. (in press). Self-assessment in personnel selection: An investigation using seasonal workers. Applied Psychology: An International Review.
- Gergen, K. J. (1984). Theory of the self: Impasse and evolution. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology (pp. 49-117). London: Academic Press.
- Ghiselli, E. E. (1956). Differentiation of individuals in terms of their predictability. Journal of Applied Psychology, 40, 374-377.
- Ghiselli, E. G. (1966). The Validity of Occupational Aptitude Tests. New York: John Wiley & Sons.
- Ghiselli, E. G. (1972). Comment on the use of moderator variables. Journal of Applied Psychology, 56, 270.

- Gimpl, M. L., & Dakin, S. R. (1984). Management and magic. California Management Review, 27, 125-136.
- Goffman, E. (1969). The Presentation on Self in Everyday Life. London: Penguin.
- Gordon, M. E., & Kleiman, L. S. (1976). The prediction of trainability using a work sample test and an aptitude test: A direct comparison. Personnel Psychology, 29, 243-253.
- Gorsuch, R. L. (1983). Factor Analysis (2nd ed.). Hillsdale, New Jersey: Lawrence Erlbaum.
- Grande, P. P. (1966). The use of self and peer ratings in a Peace Corps training programme. Vocational Guidance Quarterly, 14, 244-246.
- Grant, D. L. (1980). Issues in personnel selection. Professional Psychology, 11, 369-384.
- Grant, D. L., Katkovsky, W., & Bray, D. W. (1967). Contributions of projective techniques to assessment of management potential. Journal of Occupational Psychology, 51, 276-232.
- Graves, N. B., & Graves, T. D. (1977). Preferred adaptive strategies: An approach to understanding New Zealand's multi-cultural work force. New Zealand Journal of Industrial Relations, 6, 81-89.
- Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. American Psychologist, 35, 603-618.
- Greenwald, A. G. (1981). Self and memory. In G. H. Bower (Ed.), Psychology of Learning and Motivation (Vol. 15, pp. 201-236). New York: Academic Press.
- Greenwald, A. G., & Pratkanis, A. R. (1984). The self. In R. S. Wyer, & T. K. Srull (Eds.), Handbook of Social Cognition (Vol. 3, pp. 129-178). Hillsdale, New Jersey: Lawrence Erlbaum.
- Guilford, J. P., & Zimmerman, W. S. (1956). Fourteen dimensions of temperament. Psychological Monographs, 70(Serial No. 10).

- Guion, R. M. (1961). Criterion measurement and personnel judgments. Personnel Psychology, 19, 141-149.
- Guion, R. M. (1976). Recruiting, selection, and job placement. In M. D. Dunnette (Ed.), Handbook of Industrial and Organisational Psychology. Chicago: Rand McNally.
- Guion, R. M. (1987). Changing views for personnel selection research. Personnel Psychology, 40, 199-213.
- Guion, R. M., & Cranny, C. J. (1982). A note on concurrent and predictive validity designs: A critical reanalysis. Journal of Applied Psychology, 67, 239-244.
- Guion, R. M., & Gottier, R. F. (1965). Validity of personality measures in personnel selection. Personnel Psychology, 18, 135-164.
- Guion, R. M., & Ironson, G. H. (1983). Latent trait theory for organisational research. Organisational Behaviour and Human Performance, 31, 54-87.
- Hackett, G., & Betz, N. E. (1981). A self-efficacy approach to the career development of women. Journal of Vocational Behaviour, 18, 326-399.
- Hakel, M. D. (1985). Personnel selection and placement. Annual Review of Psychology, 37, 351-380.
- Hammer, T. H., & Landau, J. (1981). Methodological issues in the use of absence data. Journal of Applied Psychology, 66, 574-581.
- Hansford, B. C., & Hattie, J. A. (1982). The relationship between self-and achievement/performance measures. Review of Educational Research, 52, 123-142.
- Harris, J. G. (1972). Prediction of success on a distant Pacific island: Peace Corps style. Journal of Consulting and Clinical Psychology, 38, 181-190.
- Harris, R. J. (1985). A Primer of Multivariate Statistics (2nd ed.). Orlando, Florida: Academic Press.

- Hathaway, S. R., & McKinley, J. C. (1967). Minnesota Multiphasic Personality Inventory: Manual for Administration and Scoring. New York: Psychological Corporation.
- Hawk, J. (1986). Real world implications of g. Journal of Vocational Behaviour, 29, 411-414.
- Henbest, R. J., & Fehrsen, M. B. (1985). Preliminary study at the Medical University of Southern Africa on student self-assessment as a means of evaluation. Journal of Medical Education, 60, 66-68.
- Henderson, R. (1987). Investigation into staff selection practices in Canterbury businesses. Unpublished manuscript. Massey University: Palmerston North, New Zealand.
- Heneman, H. G. (1974). Comparisons of self- and superior ratings of managerial performance. Journal of Applied Psychology, 59, 638-642.
- Heneman, H. G. (1980). Self-assessment: A critical analysis. Personnel Psychology, 33, 297-300.
- Herriot, P., & Rothwell, C. (1981). Organisational choice and decision theory: Effects of employers' literature and selection interview. Journal of Occupational Psychology, 54, 17-31.
- Hesketh, B. (1974). Survey on Testing In Industry. Wellington, New Zealand: New Zealand Institute of Personnel Management.
- Hesketh, B. (1984). The changing nature of work. National Advisory Research Council Science and Technology Reviews. Wellington: New Zealand Government Printer.
- Hicks, J. (1984). Unemployment in New Zealand: A background to the problem (Massey University Economic Paper No. C8409). Palmerston North, New Zealand: Massey University.
- Hill, J. M. M., & Trist, E. L. (1953). A consideration of industrial accidents as a means of withdrawal from the work situation. Human Relations, 6, 357-380.

- Hines, G. H. (1972). Management attitudes toward industrial psychologists: A cross-cultural study. American Psychologist, 7, 123-130.
- Hines, G. H. (1973). An overview of business and society. In G. H. Hines (Ed.), Business and New Zealand Society (pp. 1-6). Wellington, New Zealand: Hicks, Smith & Sons Ltd.
- Hines, G. H. (1974). Sociocultural influences on employee expectancy and participative management. Academy of Management Journal, 17, 334-339.
- Hines, G. H. (1976). Cultural influences on work motivation. In P. Warr (Ed.), Personal Goals and Work Design (pp. 23-41). London: John Wiley & Sons.
- Hinrichs, J. R., & Haanpera, S. (1976). Reliability of measurement in situational exercises. Personnel Psychology, 28, 31-40.
- Hobert, R., & Dunnette, M. D. (1967). Development of moderator variables to enhance the prediction of managerial effectiveness. Journal of Applied Psychology, 51, 50-64.
- Hoelter, J. W. (1985). The structure of self-conception: Conceptualisation and measurement. Journal Of Applied Psychology, 49, 1392-1407.
- Hoffman, P. J., Festinger, L., & Lawrence, D. H. (1954). Tendencies toward group comparability in competitive bargaining. Human Relations, 7, 141-159.
- Hofstede, G. (1983). The cultural relativity of organisational practices and theories. Journal of International Business Studies, 13, 75-89.
- Hofstede, G. (1984). Culture's Consequences (abr. ed.). Newbury Park, California: Sage.
- Hogan, J., & Quigley, A. M. (1986). Physical standards for employment and the courts. American Psychologist, 41, 1193-1217.
- Hollenbeck, J. R., & Williams, C. R. (1986). Turnover functionality versus turnover frequency: A note on work attitudes and organisational effectiveness. Journal of Applied Psychology, 71, 606-611.

Holzbach, R. I. (1978). Rater bias in performance ratings: Superior, self-, and peer ratings. Journal of Applied Psychology, 63, 579-588.

Hough, L. M. (1984). Development and evaluation of the "Accomplishment Record" method of promoting individuals. Journal of Applied Psychology, 69, 135-146.

Hough, L. M., Keyes, M. A. & Dunnette, M. D. (1983). An evaluation of three alternative selection procedures. Personnel Psychology, 36, 261-276.

Hoult, P. P., & Smith, M. C. (1978). Age and sex differences in the number and variety of vocational choices, preferences, and aspirations. Journal of Occupational Psychology, 51, 119-125.

Hull, C. H., & Nie, N. H. (1981). SPSS Update 7-9. New York: McGraw-Hill.

Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. Journal of Vocational Behaviour, 29, 340-362.

Hunter, J. E., & Hunter R. F. (1984). Validity and utility of alternative predictors of job performance. Psychological Bulletin, 69, 72-98.

Hunter J. E., & Schmidt, F. L. (1982a). Fitting people to jobs: The impact of personnel selection on national productivity. In M. D. Dunnette & E. A. Fleishman (Eds.) Human Performance and Productivity: Vol. 1. Human Capability Assessment (pp. 233-284). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Hunter J. E., & Schmidt, F. L. (1982b). Ability tests: Economic benefits versus the issue of fairness. Industrial Relations, 21, 293-308.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). Meta-analysis: Cumulating Research Finding Across Studies. Beverly Hills, California: Sage.

Huse, E. F., & Taylor, E. K. (1962). The reliability of absence measures. Journal of Applied Psychology, 46, 159-160.

- Hutchinson, V. (1984). Alternative employment contracts. In G. Shouksmith & B. Hesketh (Eds.), Planning and training for work and leisure (Occasional Papers in Psychology, No. 11, pp. 65-74). Palmerston North New Zealand: Massey University, Psychology Department.
- Inkson, K. (1984, March). Applying Psychology in New Zealand Business. New Zealand Psychological Society Occupational Division Newsletter, pp. 2-3.
- Inkson, K. (1987). Organisational behaviour: A review of New Zealand research. New Zealand Psychologist, 16, 9-27.
- Inkson, K., Henshall, B., Marsh, N., & Ellis, G. (1986). Theory K: The Key to Excellence in New Zealand Management. Auckland, New Zealand: David Bateman Ltd.
- James, L. R., & Brett, J. M. (1984). Mediators, moderators, and tests for mediation. Journal of Applied Psychology, 69, 307-321.
- Jenkins, P. (1985, June). NZ Steel: An apprentice selection case study. Management, pp. 63-66.
- Johns, G., & Nicholson, N. (1982). The meanings of absence: New strategies for theory and research. In B. Staw and L. Cummings (Eds.), Research in Organisational Behaviour (Vol. 4, pp. 127-172). Greenwich, Connecticut: JAI Press.
- Jones, A. (1981). Inter-rater reliability in the assessment of group exercises at a UK assessment centre. Journal of Occupational Psychology, 54, 79-86.
- Jones, A., & Harrison, E. (1982). Prediction of performance in initial officer training using reference reports. Journal of Occupational Psychology, 55, 35-42.
- Jones, E. E., & Baumeister, R. (1976). The self-monitor looks at the ingratiation. Journal of Personality, 44, 654-674.
- Jones, E. E., Gergen, K. L., & Jones, R. G. (1963). Tactics of ingratiation among leaders and subordinates in a status hierarchy. Psychological Monographs, 77, (3, Whole No. 566).

- Jones, E. E., & Pittman, T. S. (1982). Towards a general theory of strategic self-presentation. In J. Suls (Ed.), Psychological Perspectives on the Self (Vol. 1, pp. 231-262). Hillsdale, New Jersey: Lawrence Erlbaum
- Jones, E. E., & Wortman, C. (1973). Ingratiation: An Attributional Approach. Morristown, New Jersey: General Learning Press.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. Psychological Review, 80, 237-251.
- Keenan, A. (1978). The selection interview: Candidates' reactions and interviewers' judgments. British Journal of Social and Clinical Psychology, 17, 201-209.
- Keenan, A., & Logue, C. (1985, April). Interviewing women. Personnel Management, p. 59.
- Keller, R. T. (1983). Predicting absenteeism from prior absenteeism, attitudinal factors, and nonattitudinal factors. Journal of Applied Psychology, 68, 536-540.
- Kelly, E. L., & Fiske, D. W. (1951). The Prediction of Performance in Clinical Psychology. Ann Arbor: University of Michigan Press.
- Kesselman, G. A., Lopez, F. M., & Lopez, F. E. (1982). The development and validation of a self-report scored in-basket test in an assessment centre setting. Public Personnel Management, 11, 228-238.
- Kihlstrom, J. F., Cantor, N. (1984). Mental representations of the self. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology (pp. 2-48). London: Academic Press.
- Kim, J., & Kohout, F. J. (1975). Multiple regression analysis: Subprogramme regression. In N. H. Nie, C. H. Hull, J. G. Jenkins, K. Steinbrenner, & D. H. Bent (Eds.), Statistical Package for the Social Sciences (2nd ed., pp. 320-367). New York: McGraw-Hill.
- King, M. R., & Manaster, G. J. (1977). Body image, self-esteem, expectations, self-assessments, and actual success in a simulated job interview. Journal of Applied Psychology, 62, 589-594.

- Kinicki, A. J., & Lockwood, C. A. (1985). The interview process: An examination of factors recruiters use in evaluating job applicants. Journal of Vocational Behaviour, 26, 117-125.
- Kinslinger, H. J. (1966). Application of projective techniques in personnel psychology since 1940. Psychological Bulletin, 66, 134-149.
- Kirchner, W. K. (1965). Relationships between supervisory and subordinate ratings for technical personnel. Journal of Industrial Psychology, 3, 57-60.
- Klecka, W. R. (1975). Discriminant analysis. In N. H. Nie, C. H. Hull, J. G. Jenkins, K. Steinbrenner, & D. H. Bent (Eds.), Statistical Package for the Social Sciences (2nd ed., pp. 434-467). New York: McGraw-Hill.
- Klecka, W. R. (1984). Discriminant Analysis. Sage University Paper Series Quantitative Applications in the Social Sciences. Beverley Hills and London: Sage Publications.
- Klieman, L. S., & Faley, R. H. (1985). Implications of professional and legal guidelines for court decisions involving criterion related validity: A review and analysis. Personnel Psychology, 38, 803-833.
- Klimoski, R., & Brickner, M. (1987). Why do assessment centres work? The puzzle of assessment centre validity. Personnel Psychology, 40, 243-260.
- Klimoski, R. J., & London, M. (1974). Role of the rater in performance appraisal. Journal of Applied Psychology, 59, 445-451.
- Klimosky, R. J., & Strickland, W. J. (1977). Assessment centres: Valid or merely prescient. Personnel Psychology, 30, 353-361.
- Knouse, S. B. (1983). The letter of recommendation: Specificity and favourability of information. Personnel Psychology, 36, 331-341.
- Kooker, E. W. (1974). Changes in ability of graduate students in education to assess their own test performance as related to their Miller Analogies scores. Psychological Reports, 35, 97-98.

- Korman, A. K. (1966). Self-esteem variable in vocational choice. Journal of Applied Psychology, 50, 479-486.
- Korman, A. K. (1967). Self-esteem as a moderator of the relationship between self-perceived abilities and vocational choice. Journal of Applied Psychology, 51, 65-67.
- Korman, A. K. (1970). Toward a hypothesis of work behaviour. Journal of Applied Psychology, 54, 31-51.
- Kriedt, P. H., & Dawson, R. I. (1961). Response set and the prediction of clerical job performance. Journal of Applied Psychology, 45, 175-178.
- Krumboltz, J. D., Mitchell, A. M., & Jones, G. B. A social learning theory of career selection. The Counselling Psychologist, 6, 71-81.
- Kulka, R. A. (1982). Idiosyncrasy and circumstance: Choices and constraints in the research process. In J. E. McGrath, J. Martin, & R. A. Kulka, (Eds.), Judgment Calls in Research (pp. 41-68). Beverly Hills, California: Sage Publications.
- Landy, F. J. (1986). Stamp collecting versus science: Validation as hypothesis testing. American Psychologist, 41, 1183-1192.
- Landy, F. J., & Farr, J. L. (1983). The Measurement of Work Performance. New York, Academic Press.
- Landy, F. J., Farr, J. L., & Jacobs, R. R. (1982). Utility concepts in performance measurement. Organisational Behaviour and Human Performance, 30, 15-40.
- Latham, G. P., & Saari, L. M. (1984). Do people do what they say? Further studies on the situational interview. Journal of Applied Psychology, 69, 469-573.
- Latham, G. P., Saari, L. M., Pursell, E. D., & Campion, M. A. (1980). The situational interview. Journal of Applied Psychology, 69, 422-427.

- Lawler, E. E. (1967). The multitrait-multirater approach to measuring managerial job performance. Journal of Applied Psychology, 51, 369-381.
- Lefkowitz, J., & Katz, M. L. (1969). Validity of exit interviews. Personnel Psychology, 22, 445-455.
- Levine, E. L. (1980). Introductory remarks for the symposium "Organisational applications of self-appraisal and self-assessment: Another look". Personnel Psychology, 33, 259-262.
- Levine, E. L., Edgell, S. L., & Ash, R. A. (1978, March). Validity of a matching approach to personnel selection. Paper presented at the meeting of the Southeastern Psychological Association, Atlanta, Georgia, March, 1978.
- Levine, E. L., Flory, A., & Ash, A. A. (1977). Self-assessment in personnel selection. Journal of Applied Psychology, 62, 428-435.
- Lewis, C. (1984, January). What's new in selection? Personnel Management, pp. 14-16.
- Liden, R. C., & Parsons, C. K. (1986). A field study of job applicant interview perceptions, alternative opportunities, and demographic characteristics. Personnel Psychology, 39, 109-122.
- Linn, R. L. (1984). Selection bias: Multiple meanings. Journal of Educational Measurement, 21, 33-47.
- Lunneborg, C. E. (1982). Systematic biases in brief self-ratings of vocational qualifications. Journal of Vocational Behaviour, 20, 255-275.
- Mabe, P. A., & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. Journal of Applied Psychology, 67, 280-296.
- Maccoby, E. E., & Jacklin, C. (1974). The Psychology of Sex Differences. Stanford, California: Stanford University Press.

- MacKay, C., Cox, T., Burrows, G., & Lazzerini, T. (1978). An inventory for the measurement of self-reported stress and arousal. British Journal of Social and Clinical Psychology, 17, 283-284.
- Macy, B. A., & Mirvis, P. H. (1983). Assessing rates and costs of individual work behaviours. In S. E. Seashore, E. E. Lawler, P. H. Mirvis, & C. Cammann (Eds.), Assessing Organisational Change (pp. 139-177). New York: Wiley.
- Makin, P. J., & Robertson, I. T. (1986, November). Selecting the best selection techniques. Personnel Management, 38-40.
- Markus, H. (1977). Self-schemata and processing information about the self. Journal of Personality and Social Psychology, 35, 63-78.
- Markus, H. (1980). The self in thought and memory. In D. M. Wegner & R. R. Vallacher (Eds.), The Self in Social Psychology. New York: Oxford University Press.
- Markus, H. (1983). Self-knowledge: An expanded view. Journal of Personality, 51, 544-565.
- Markus, H., & Sentis, K. (1982). The self in social information processing. In J. Suls (Ed.), Psychological Perspectives on the Self. Hillsdale, New Jersey: Erlbaum.
- Markus, H., & Wurf, E. (1987). The dynamic self-concept. Annual Review of Psychology, 38, 299-337.
- Marlowe, D., & Crowne, D. P. (1964). Social desirability and response to perceived situational demands. Journal of Consulting Psychology, 25, 109-115.
- Marsh, H. W. (1986). Self-serving (bias?) in academic attributions: Relation to academic achievement and self-concept. Journal of Educational Psychology, 78, 190-200.
- Marsh, H. W., Barnes, J., Cairnes, L., & Tidman, M. (1984). Self-Description Questionnaire: Age effects in the structure and level of self-concept for preadolescent children. Journal of Educational Psychology, 76, 940-956.

- Marsh, H. W., & Shavelson, R. J. (1985). Self-concept: Its multifaceted hierarchical structure. Educational Psychologist, 20, 107-125.
- Marsh, H. W., Smith, I. D., & Barnes, J. (1985). Multi-dimensional self-concepts: Relations with sex and academic achievement. Journal of Educational Psychology, 77, 581-596.
- Martin, J. (1982). A garbage can model of the research process. In J. E. McGrath, J. Martin, & R. A. Kulka, (Eds.), Judgment Calls in Research (pp. 17-39). Beverly Hills, California: Sage Publications.
- Mason, B. L., & Cole, B. H. (1977). The Meremere coal gang: A study of a multicultural work gang. New Zealand Journal of Industrial Relations, 6, 90-97.
- Mathews, R. M., & Fawcett, S. B. (1984). Building the capacities of job candidates through behavioural instruction. Journal of Community Psychology, 12, 123-129.
- Matteson, M. T. (1978). An alternative approach to using biographical data for predicting job success. Journal of Occupational Psychology, 51, 155-162.
- Mayfield, E. C. (1964). The selection interview: A re-evaluation of published research. Personnel Psychology, 17, 239-260.
- Mayo, G. D., & Manning, W. H. (1961). Motivation measurement. Educational and Psychological Measurement, 21, 73-83.
- McClelland, D. C. (1972). Opinions predict opinions: So what else is new? Journal of Consulting and Clinical Psychology, 38, 325-326.
- McCormick, I. A, Walkley, F. H., & Green, D. E. (1986). Comparative perceptions of driver ability: A confirmation and expansion. Accident Analysis and Prevention, 18, 205-208.
- McEnery, J., & McEnery, J. M. (1987). Self-rating in management training needs assessment: A neglected opportunity? Journal of Occupational Psychology, 60, 49-60.
- McEvoy, G. M., & Cascio, W. F. (1985). Strategies for reducing employee turnover: A meta-analysis. Journal of Applied Psychology, 70, 342-353.

- McGrath, J. E., Martin, J., & Kulka, R. A. (1982). Judgment Calls in Research. Beverly Hills, California: Sage Publications.
- McGuire, W. J. (1984). Search for the self: Going beyond self-esteem and the reactive self. In R. A. Zucker, J. Aronoff, & A. I. Rabin (Eds.), Personality and the Prediction of Behaviour (pp. 73-120). Orlando, Florida: Academic Press.
- McGuire, W. J., & McGuire, C. V. (1982). Significant others in self-space: Sex differences and developmental trends in the social self. In J. Suls (Ed.), Psychological Perspectives on the Self (Vol.1, pp. 71-96). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- McLennan, R., Inkson, K., Dakin, S., Dewe, P., & Elkin, G. (1987). People and Enterprises: Human Behaviour in New Zealand Organisations. Sydney: Holt, Rinehart, and Winston.
- Mead, G. H. (1934). Mind, Self, and Society. Chicago: Chicago University Press.
- Mergagee, E. I. (1972). The California Psychological Inventory Handbook. San Francisco: Jossey-Bass.
- Merritt-Haston, R., & Wexley, K. N. (1983). Educational requirements: Legality and Validity. Personnel Psychology, *36*, 743-753.
- Meyer, H. H. (1980). Self-appraisal of job performance. Personnel Psychology, *33*, 291-295.
- Meyer, H. H. (1987). Predicting supervisory ratings versus promotional progress in test validation studies. Journal of Applied Psychology, *72*, 696-697.
- Mihal, W. L., & Graumenz, J. L. (1984). An assessment of the accuracy of self-assessment for career decision making. Journal of Vocational Behaviour, *25*, 245-253.
- Miner, J. B. (1977). Motivation to Manage. Atlanta: Organisational Measurement Systems Press.
- Minor, S. W., & Roberts, A. M. (1984). Effects of congruence between locus of control and task instructions on self-efficacy. Psychological Reports, *54*, 39-42.

- Virfin, K., & Hesketh, B. (1985, January). What do occupational psychologists see as their future roles. New Zealand Psychological Society Occupational Division Newsletter, pp. 4-9.
- Mischel, W. (1965). Predicting the success of Peace Corps volunteers in Nigeria. Journal of Personality and Social Psychology, 1, 510-517.
- Mischel, W. (1968). Personality and Assessment. New York: Wiley.
- Mischel, W. (1972). Direct versus indirect personality assessment: Evidence and implications. Journal of Consulting and Clinical Psychology, 38, 319-324.
- Mischel, W. (1984). On the predictability of behaviour and the structure of personality. In R. A. Zucker, J. Aronoff, & A. I. Rabin (Eds.), Personality and the Prediction of Behaviour (pp. 269-305). Orlando, Florida: Academic Press.
- Mischel, W., & Mischel, H. N. (1977). Self-control and the self. In T. Mischel (Ed.), The Self: Psychological and Philosophical Issues (pp. 31-64). Oxford: Blackwell.
- Mitchell, J. V. (1985). The Ninth Mental Measurements Yearbook (Vol. 2). Lincoln, Nebraska: University of Nebraska Press.
- Monahan, C. J., & Muchinsky, P. M. (1983). Three decades of personnel selection research: A state-of-the-art analysis and evaluation. Journal of Occupational Psychology, 56, 215-225.
- Moreland, R., Miller, J., & Laucka, R. (1981). Academic achievement and self-evaluations of academic performance. Journal of Educational Psychology, 73, 335-334.
- Morse, S. J., & Gergen, K. J. (1970). Social comparison, self-consistency and the concept of self. Journal of Personality and Social Psychology, 16, 148-156.
- Moskowitz, D. S. (1986). Comparison of self-reports, reports by knowledgeable informants, and behavioural observation data. Journal of Personality, 54, 294-317.

- Mount, M. K. (1984). Psychometric properties of subordinate ratings of managerial performance. Personnel Psychology, 37, 687-702.
- Mowday, R. T., Porter, L. W., & Steers, R. M. (1982). Employee-Organisation Linkages: The Psychology of Commitment, Absenteeism and Turnover. New York, Academic Press.
- Muchinsky, P. M. (1977). Employee absenteeism: A review of the literature. Journal of Vocational Behaviour, 10, 316-340.
- Muchinsky, P. M. (1979). The use of reference reports in personnel selection: A review and evaluation. Journal of Occupational Psychology, 52, 287-297.
- Muchinsky, P. M., & Morrow, P. C. (1980). A multidisciplinary model of voluntary employee turnover. Journal of Vocational Behaviour, 17, 263-290.
- Muchinsky, P. M., & Tuttle, M. L. (1979). Employee turnover: An empirical and methodological assessment. Journal of Vocational Behaviour, 14, 43-77.
- Munchus, G. (1986). Polygraph use in business and industry: A human resource management analysis. Managerial Psychology, 1, 31-35.
- Murphy, K. R. (1986). When your top choice turns you down: Effect of rejected offers on the utility of selection tests. Psychological Bulletin, 99, 133-138.
- Nathanson, S. H. (1975). Evaluation and analysis of the Cleff Job Matching System, Final Report (Contract No. B2C 5415). Newport Beach, California, Ultrasystems.
- Nedd, A. N. B., & Marsh, N. R. (1979). Social traditionalism and personality: An empirical investigation of the inter-relationships between social values and personality attributes. International Journal of Psychology, 14, 73-82.
- Nedd, A. N. B., & Marsh, N. R. (1980). A cross-cultural test of the personality integration hypothesis. Journal of Personality, 48, 293-305.
- Negandhi, A. R. (1983). Cross-cultural management research: Trend and future directions. Journal of International Business Studies, 13, 17-28.

- Neidig, R. D., & Neidig, P. J. (1984). Multiple assessment centre exercises and job relatedness. Journal of Applied Psychology, 69, 182-186.
- Veiner, A. G., & Owens, W. A. (1985). Using biodata to predict job choice among college graduates. Journal of Applied Psychology, 70, 127-136.
- New Zealand Institute of Personnel Management. (1979). Personnel Management 1978. Wellington, New Zealand: Author.
- Nicholson, N., & Goodge, P. M. (1976). The influence of social, organisational, and biographical factors on female absence. Journal of Management Studies, 13, 234-254.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84, 231-259.
- Noe, R. A., & Steffy, B. D. (1987). The influence of individual characteristics and assessment centre evaluation on career exploration behaviour and job involvement. Journal of Vocational Behaviour, 30, 187-202.
- Norusis, M. J. (1985). SPSS^X Advanced Statistics Guide. New York: McGraw-Hill.
- Olian, J. D. (1984). Genetic screening for employment purposes. Personnel Psychology, 37, 423-438.
- Owens, W. A. (1971). A quasi-actuarial basis for individual assessment. American Psychologist, 26, 992-999.
- Owens, W. A. (1976). Background data. In Dunnette, M. D. (Ed.), Handbook of Industrial and Organisational Psychology. Chicago: Rand McNally.
- Owens, W. A., & Schoenfeldt, L. F. (1979). Towards a classification of persons. Journal of Applied Psychology, 65, 569-607.
- Pannone, R. D. (1984). Predicting test performance: A content valid approach to screening applicants. Personnel Psychology, 37, 507-514.

- Parisher, D., Rios, B., & Reilley, R. R. (1979, February). Psychologists and psychological services in urban police departments. Professional Psychology, pp. 6-7.
- Parker, J. W., Taylor, E. K., Barrett, R. S., & Martens, (1959). Rating scale content: III. Relationship between supervisory and self-ratings. Personnel Psychology, 12, 49-63.
- Parsons, C. K., Herold, D. M., & Leatherwood, M. L. (1985). Turnover during initial employment: A longitudinal study of the role of causal attributions. Journal of Applied Psychology, 70, 337-341.
- Pearlman, K., Schmidt, F. L., & Hunter, J. E. (1980). Validity generalisation results for tests used to predict job proficiency and training success in clerical occupations. Journal of Applied Psychology, 65, 373-406.
- Pervin, L. A. (1976). A free-response description approach to the analysis of person-situation interaction. Journal of Personality and Social Psychology, 34, 465-474.
- Peters, W. S., & Champoux, J. E. (1979). The role and analysis of moderators in organisational research. In R. T. Mowday & R. M. Steers, (Eds.), Research in Organisations: Issues and controversies. Santa Monica, California: Goodyear.
- Peterson, D. R. (1965). Scope and generality of verbally defined personality factors. Psychological Review, 72, 48-59.
- Phipps, P. R., & Toulson, P. K. (1982, August). Scientific or Intelligible: The Occupational Psychologist's Dilemma. Paper presented at the Annual Conference of the New Zealand Psychological Society. Hamilton, New Zealand.
- Pirikahu, N., MacPherson, J., Gibbs, L., Kohu, S., & Ponter, B. (1984). The Patea experience: Government aid and emergent industries in a rural area with high unemployment. In G. Shouksmith & B. Hesketh (Eds.), Planning and training for work and leisure (Occasional Papers in Psychology, No. 11, pp. 48-61). Palmerston North New Zealand: Massey University, Psychology Department.

- Plorde, D. S., Wollitzer, A. O., & Blossom, H. J. (1985). A self-assessment questionnaire for patient educators. Journal of Medical Education, 60, 728-730.
- Popovich, P., & Wanous, J. P. (1982). The realistic job preview as a persuasive communication. Academy of Management Review, 7, 570-578.
- Porter, L. W., & Steers, R. M. (1973). Organisational, work and personal factors in turnover and absenteeism. Psychological Bulletin, 80, 151-186.
- Powell, G. N. (1984). Effects of job attributes and recruiting practices on applicant decisions: A comparison. Personnel Psychology, 37, 721-732.
- Premack, S. L., & Wanous, J. P. (1985). A meta-analysis of realistic job preview experiments. Journal of Applied Psychology, 70, 706-719.
- Prien, E. P., & Liske, R. E. (1962). Assessments of higher level personnel: III. Rating criteria: A comparative analysis of supervisor ratings and incumbent self-ratings of job performance. Personnel Psychology, 15, 187-194.
- Primoff, E. S. (1980). The use of self-assessments in examining. Personnel Psychology, 33, 283-290.
- Prior, J. B., Gibbons, F. X., Wicklund, R. A., Fazio, R. H., & Hood, R. (1977). Self-focused attention and self-report validity. Journal of Personality, 45, 513-527.
- Pym, D. L. A., & Auld, H. D. (1965). The self-rating as a measure of employee satisfactoriness. Occupational Psychology, 39, 103-113.
- Quattrone, G. A., & Jones, E. E. (1978). Selective self-disclosure with and without correspondent performance. Journal of Experimental and Social Psychology, 14, 511-526.
- Rafaelli, A., & Klimoski, R. J. (1983). Predicting sales success through handwriting analysis: An evaluation of the effects of training and handwriting sample content. Journal of Applied Psychology, 68, 212-217.

- Raju, N. S., & Burke, M. J. (1983). Two new procedures for studying validity generalisation. Journal of Applied Psychology, 68, 382-395.
- Raven, J. C. (1965). Advanced Progressive Matrices: Sets I and II. London: William Grieve & Sons.
- Raza, S. M., & Carpenter, B. N. (1987). A model of hiring decisions in real employment interviews. Journal of Applied Psychology, 72, 596-603.
- Reilly, R. R., & Chao, G. T. (1982). Validity and fairness of some alternative employee selection procedures. Personnel Psychology, 35, 1-62.
- Rhodes, S. R. (1983). Age-related differences in work attitudes and behaviour: A review and conceptual analysis. Psychological Bulletin, 93, 328-367.
- Ritchie, J. E. (1973). The Employment of Minority Groups in New Zealand Society. Wellington, New Zealand: Hicks Smith.
- Ritchie, R. J., & Moses, J. L. (1983). Assessment centre correlates of women's advancement into middle management: A 7-year longitudinal analysis. Journal of Applied Psychology, 68, 227-231.
- Robertson, I. T., & Downs, S. (1979). Learning and the prediction of performance: Development of Trainability Testing in the United Kingdom. Journal of Applied Psychology, 64, 42-50.
- Robertson, I. T., Gratton, L., & Sharpley, D. (1987). The psychometric properties and design of managerial assessment centres: Dimensions into exercises won't go. Journal of Occupational Psychology, 60, 187-195.
- Robertson, I. T., & Kandola, R. S. (1982). Work sample tests: Validity, adverse impact, and applicant reaction. Journal of Occupational Psychology, 55, 171-183.
- Robertson, I. T., & Makin, P. J. (1986). Management selection in Britain: A survey and critique. Journal of Occupational Psychology, 59, 45-57.

- Robinson, J. P., & Shaver, P. R. (1973). Measures of Social Psychological Attitudes (Rev. Ed.). Michigan: The Institute for Social Research, University of Michigan.
- Roe, A. (1953). A psychological study of eminent psychologists and anthropologists, and a comparison with biological and physical scientists. Psychological Monographs, 67, (Whole No. 352).
- Rogers, T. B. (1981). A model of the self as an aspect of the human information processing system. In N. Cantor, & J. Kihlstrom (Eds.), Personality, Cognition, and Social Interaction (pp 193-214). Hillsdale, New Jersey: Lawrence Erlbaum.
- Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. Journal of Personality and Social Psychology, 35, 677-688.
- Rotter, J. B. (1966). Generalised expectancies for internal versus external control of reinforcement. Psychological Monographs, 80(1, Whole No. 609).
- Russell, G. J. (1985). Individual decision processes in an assessment centre. Journal of Applied Psychology, 70, 737-746.
- Ryan, A. M., & Sackett, P. R. (1987). A survey of individual assessment practices by I/O psychologists. Personnel Psychology, 40, 455-488.
- Rynes, S. L., & Miller, H. E. (1983). Recruiter and job influences on candidates for employment. Journal of Applied Psychology, 68, 147-154.
- Sackett, P. R., & Decker, P. J. (1979). Detection of deception in the employment context: A review and critical analysis. Personnel Psychology, 32, 487-506.
- Sackett, P. R., & Dreher, G. F. (1982). Constructs and assessment centre dimensions: Some troubling empirical findings. Journal of Applied Psychology, 67, 401-410.
- Sackett, P. R., & Dreher, G. F. (1984). Situation specificity of behaviour and assessment centre evaluation: A rejoinder to Neidig and Neidig. Journal of Applied Psychology, 69, 187-190.

- Sackett, P. R. & Harris, M. M. (1984). Honesty testing for personnel selection: A review and critique. Personnel Psychology, 37, 221-245.
- Sackett, P. R., Tenopyr, M. L., Schmitt, N., & Kehoe, J. (1985). Commentary on forty questions about validity generalisation and meta-analysis. Personnel Psychology, 38, 697-798.
- Sanders, G. S. (1985). Unbiased evaluations of performance: Artifact or research topic. British Journal of Social Psychology, 24, 159-160.
- Sanders, G. S., & Mullen, B. D. (1984). The extent of logical deduction in self-evaluation of abilities. British Journal of Social Psychology, 23, 9-19.
- Schendel, J. D., & Hagman, J. D. (1982). On sustaining procedural skills over a prolonged retention interval. Journal of Applied Psychology, 67, 605-610.
- Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalisation. Journal of Applied Psychology, 62, 529-540.
- Schmidt, F. L., & Hunter, J. E. (1980). The future of criterion related validity. Personnel Psychology, 33, 41-60.
- Schmidt, F. L., & Hunter, J. E. (1981). Employment testing: Old theories and new research findings. American Psychologist, 36, 1128-1137.
- Schmidt, F. L., Hunter, J. E., & Caplan, J. R. (1981). Validity generalisation results for two jobs in the petroleum industry. Journal of Applied Psychology, 66, 261-273.
- Schmidt, F. L., Hunter, J. E., Outerbridge, A. N., & Trattner, M. H. (1986). The economic impact of job selection methods on size, productivity, and payroll costs of the Federal work force: An empirically based demonstration. Personnel Psychology, 39, 1-30.
- Schmidt, F. L., & Hunter, J. E., & Pearlman, K. (1982). Progress in validity generalisation: Comments on Callender and Osburn and further developments. Journal of Applied Psychology, 67, 835-845.

- Schmidt, F. L., Pearlman, K., & Hunter, J. E. (1980). The validity and fairness of employment and educational tests for Hispanic Americans: A review and analysis. Personnel Psychology, 33, 705-724.
- Schmidt, F. L., Pearlman, K., Hunter, J. E., & Hirsh, H. R. (1985). Forty questions about validity generalisation and meta-analysis. Personnel Psychology, 38, 697-798.
- Schmitt, N. (1976). Social and situational determinants of interview decisions: Implications for the employment interview. Personnel Psychology, 29, 79-101.
- Schmitt, N., & Coyle, B. W. (1976). Applicant decisions in the employment interview. Journal of Applied Psychology, 61, 184-192.
- Schmitt, N., Ford, J. K., & Stults, D. M. (1986). Changes in self-perceived ability as a function of performance in an assessment centre. Journal of Occupational Psychology, 59, 327-335.
- Schmitt, N., Gooding, R. Z., Noe, R. D., & Kirsch, M. (1984). Meta-analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. Personnel Psychology, 37, 407-422.
- Schuh, A. J. (1967). The predictability of employee tenure: A review of the literature. Personnel Psychology, 20, 133-152.
- Scott, W. A., & Johnson, R. C. (1972). Comparative validities of direct and indirect personality tests. Journal of Consulting and Clinical Psychology, 38, 301-318.
- Shackleton, V., & Anderson, N. (1987). Personnel recruitment and selection. In B. M. Bass & P. J. D. Drenth (Eds.), Advances in Organisational Psychology (pp. 19-33). Newbury Park: Sage.
- Shaffer, D. R., & Tomarelli, M. (1981). Bias in the ivory tower: An unintended consequence of the Buckley Amendment for graduate admissions. Journal of Applied Psychology, 66, 7-11.

- Shaffer, G. S., Saunders, V., & Owens, W. A. (1986). Additional evidence for the accuracy of biographical data: Long-term retest and observer ratings. Personnel Psychology, 39, 791-809.
- Shapiro, G. L., & Dessler, G. (1985). Are self-appraisals more realistic among professionals or nonprofessionals in health care? Public Personnel Management, 14, 285-291.
- Shavelson, R. J., & Bolus, R. (1982). Self-concept: The interplay of theories and methods. Journal of Educational Psychology, 74, 3-17.
- Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. Review of Educational Research, 46, 407-441.
- Sherer, M., & Adams, C. H. (1983). Construct validation of the self-efficacy scale. Psychological Reports, 53, 899-902.
- Sherer, M., Maddux, J. E., Mercandante, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R. W. (1982). The self-efficacy scale: Construction and validation. Psychological Reports, 51, 398-406.
- Sherman, A. R. (1972). Real-life exposure as a primary therapeutic factor in the desensitisation treatment of fear. Journal of Abnormal Psychology, 79, 19-28.
- Shore, L. M., & Thornton, G. C. (1986). Effects of gender on self- and supervisory ratings. Academy of Management Journal, 29, 115-129.
- Shouksmith, G. (1987). Emerging personnel values in changing societies. In B. M. Bass & P. J. D. Drenth (Eds.), Advances in Organisational Psychology (pp. 19-33). Newbury Park: Sage.
- Shouksmith, G., & Petersen, D. (1982). A report on the selection of diplomatic and administrative trainees. Unpublished manuscript, Massey University, Department of Psychology, Palmerston North, New Zealand:
- Shrauger, J. S. (1975). Responses to evaluation as a function of initial self-perceptions. Psychological Bulletin, 82, 581-596.

- Shrauger, J. S., & Osberg, T. M. (1981). The relative accuracy of self-predictions and judgments by others in psychological assessment. Psychological Bulletin, 90, 232-251.
- Shrauger, J. S., & Patterson, M. B. (1974). Self-evaluation and the selection of dimensions for evaluating others. Journal of Personality, 42, 569-585.
- Silverman, W. H., Dalessio, A., Woods, S. B., & Johnson, R. L. (1986). Influence of assessment centre methods on assessors' ratings. Personnel Psychology, 39, 565-578.
- Smith, B. M. (1966). Explorations in competence: A study of Peace Corps teachers in Ghana. American Psychologist, 21, 555-566.
- Smith, E. R., & Miller, F. D. (1978). Limits on perception of cognitive processes: A reply to Nisbett and Wilson. Psychological Review, 85, 355-362.
- Smith, G. A., Dowling, P. J., & Barry, B. (1985). Australian managers' norms on the 16PF. Australian Psychologist, 20, 293-301.
- Smith, K. F., & Baldauf, J. R. (1982). The concurrent validity of self-rating with interviewer rating on the Australian Second Language Proficiency Scale. Educational and Psychological Measurement, 42, 1117-1125.
- Smith, M. C., & George, D. I. (1986a, March). Stupid personality tests: Who's conning who? Management, p. 50.
- Smith, M. C., & George, D. I. (1986b, July). Assessment centres: A test and taste for job applicants. Management, pp. 26-27.
- Smith, M. C., & George, D. I. (1987). Weighted application forms for personnel selection: A comparison of old and new methodologies. Australian Psychologist, 22, 361-375.
- Smith, P. C. (1976). Behaviour, results and organisational effectiveness: The problem of criteria. In Dunnette, M. D. (Ed.), Handbook of Industrial and Organisational Psychology (pp. 745-776). Chicago: Rand McNally.

- Snyder, R. A., & Bruning, N. S. (1979). Sex differences in perceived competence: An across organisations study. Administration in Social Work, 3, 349-358.
- Snyder, M., & Skrypnek, B. J. (1981). Testing hypotheses about the self: Assessments of job suitability. Journal of Personality, 49, 193-211.
- South, J. C. (1980). Fakability and the Engineer Performance Description Form. Personnel Psychology, 33, 371-375.
- Spoonley, P. (1978). The multicultural work force: The role of employers as gatekeepers. New Zealand Journal of Industrial Relations, 3, 63-70.
- Springbett, B. M. (1958). Factors affecting the final decision in the employment interview. Canadian Journal of Psychology, 12, 13-22.
- SPSS Inc. (1986). SPSS^X User's Guide (2nd ed.). New York: McGraw-Hill.
- SRA (Science Research Associates, Inc) (1954). Technical Supplement. Chicago: Author.
- Staley, C. C., & Shockley-Zalaback, P. (1986). Communication proficiency and future training needs of the female professional: Self-assessment vs., supervisors' evaluations. Human Relations, 39, 891-202.
- Steel, R. P., & Ovalle, N. K. (1984). Self-appraisal based upon supervisory feedback. Personnel Psychology, 37, 667-685.
- Steers, R. M., & Mowday, R. (1981). Employee turnover and post-decision accommodation processes. In L. Cummings & B. Staw (Eds.), Research in Organisational Behaviour (Vol. 3, pp. 235-281). Greenwich, Connecticut: JAI Press.
- Steers, R. M., & Rhodes, S. R. (1978). Major influences on employee attendance: A process model. Journal of Applied Psychology, 63, 391-407.
- Steger, J. A., Simmons, W. L., & Lavelle, S. (1973). Accuracy of prediction of own performance as a function of locus of control. Psychological Reports, 33, 59-62.

- Stumpf, S. A., Austin, E. J., & Hartman, K. (1984). The impact of career exploration and interview readiness on interview performance and outcomes. Journal of Vocational Behaviour, 24, 221-235.
- Sudman, S., & Bradburn, N. M. (1983). Asking Questions: A Practical Guide to Questionnaire Design. San Francisco: Jossey-Bass.
- Suls, J. & Mullen, B. (1982). From the cradle to the grave: Comparison and self-evaluation across the lifespan. In J. Suls (Ed.), Psychological Perspectives on the Self (Vol. 1, pp. 97-125). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Tabachnick, B. G., & Fidell, L. S. (1983). Using Multivariate Statistics. New York: Harper Row.
- Taft, R. (1948). Use of the "group situation observation method" in the selection of trainee executives. Journal of Applied Psychology, 32, 587-594.
- Task Force on Assessment Centre Standards. (1980). Standards and ethical considerations for assessment centre operations. The Personnel Administrator, 25, 35-38.
- Taylor, A. J. W., & McCormick, I. A. (1985). The prediction of performance on the International Biomedical Expedition to the Antarctic (IBEA). Unpublished manuscript, Victoria University, Psychology Department, Wellington, New Zealand.
- Taylor, E., & Nevis, E. (1957). The use of projective techniques in management selection. In M. Dooher & E. Marting (Eds.), Selection of Management Personnel (Vol. 1). New York: American Management Association.
- Taylor, K. M., & Betz, N. E. (1983). Applications of self-efficacy theory to the understanding and treatment of career indecision. Journal of Vocational Behaviour, 22, 63-81.
- Taylor, S. M., & Sniezek, J. A. (1984). The college recruitment interview: Topical content and applicant reactions. Journal of Occupational Psychology, 57, 157-168.

- Telenson, P. A., Alexander, R. A., & Barrett, G. V. (1983). Scoring the biographical information blank: A comparison of three weighting techniques. Applied Psychological Measurement, 7, 73-80.
- Tenopyr, M. L., & Oeltjen, P. D. (1982). Personnel selection and classification. Annual Review of Psychology, 33, 581-681.
- Tesser, A., & Campbell, J. (1983). Self-definition and self-evaluation maintenance. In J. Suls & A. G. Greenwald (Eds.), Psychological Perspectives on the Self (Vol. 1, pp. 1-31). Hillsdale, New Jersey: Lawrence Erlbaum.
- Tharenou, P., & Harker, P. (1984). Moderating influence of self-esteem on relationships between job complexity, performance, and satisfaction. Journal of Applied Psychology, 69, 623-622.
- Thompson, D. E., & Thompson, T. A. (1982). Court standards for job analysis in test validation. Personnel Psychology, 35, 865-874.
- Thornton, G. C. (1968). The relationship between supervisory and self-appraisals of executive performance. Personnel Psychology, 21, 441-445.
- Thornton, G. C. (1980). Psychometric properties of self-appraisals of job performance. Personnel Psychology, 33, 263-271.
- Thornton, G. C., & Byham, W. C. (1982). Assessment Centres and Managerial Performance. New York: Academic Press.
- Toh Eng Lim. (1981). Assessment centres in New Zealand. Unpublished masters thesis, Massey University, Palmerston North, New Zealand.
- Touhey, J. C. (1971). A symbolic interactionist approach to self-referent behaviour. Psychological Reports, 29, 87-90.
- Touhey, J. C. (1972). Studies in symbolic interaction: II. An experimental analysis of self-referent behaviour. The Psychological Record, 22, 325-331.

- Toulson, P. K. (1983, August). Lessons from the application of psychology in the armed forces. In K. Inkson (Chair), Psychology in Management. Symposium conducted at the Annual Conference of the New Zealand Psychological Society. Auckland, New Zealand.
- Toulson, P. K., & Williams, W. C. J. (1979). History of personnel research and psychological services in the New Zealand armed forces (DPU Technical Note 5/79). Wellington, New Zealand: Defence Headquarters, Defence Psychology Unit.
- Townsley, G. C. (1986, April). To 16PF or not to 16PF [letter to the editor]. Management, p. 76.
- Trope, Y. (1983). Self-assessment in achievement behaviour. In J. Suls & A. G. Greenwald (Eds.), Psychological Perspectives on the Self (Vol. 2, pp. 93-122). Hillsdale, New Jersey: Lawrence Erlbaum.
- Tullar, W. L., & Barrett, G. V. (1976). The future autobiography as a predictor of sales success. Journal of Applied Psychology, 61, 371-373.
- Turnage, J. J., & Muchinsky, P. M. (1984). A comparison of the predictive validity of assessment centre evaluations versus traditional measures in forecasting supervisory job performance: Interpretive implications of criterion distortion for the assessment paradigm. Journal of Applied Psychology, 69, 595-602.
- Twinn, W. J. (1977). The Kiwi and the work ethic. New Zealand Journal of Industrial Relations, 2, 97-100.
- Tyler, B., & Miller, K. (1986). The use of tests by psychologists: Report on a survey by BPS members. Bulletin of the British Psychological Society, 39, 405-410.
- Ulrich, L., & Trumbo, D. (1965). The selection interview since 1949. Psychological Bulletin, 63, 100-116.
- Vallacher, R. R. (1980). An introduction to self theory. In D. M. Wegner & R. R. Vallacher (Eds.), The Self in Social Psychology (pp. 3-30). New York: Oxford University Press.

- Vries, N. K., & van Kippenberg, A. (1985). Logical deduction and bias in self-evaluation of abilities: A reply to Sanders and Mullen. British Journal of Social Psychology, 24, 157-158.
- Wagner, R. (1949). The employment interview: A critical summary. Personnel Psychology, 2, 17-46.
- Wallis, D. (1980). Sex discrimination and the law. Bulletin of the British Psychological Society, 33, 1-5.
- Watson, C. J. (1981). An evaluation of some aspects of the Steers and Rhodes model of employee attendance. Journal of Applied Psychology, 66, 385-389.
- Watson, D. (1982). The actor and the observer: How are their perceptions of causality different? Psychological Bulletin, 92, 682-700.
- Watson, G., & Glaser, E. M. (1964). Watson-Glaser Critical Thinking Appraisal. New York: Harcourt, Brace and World Inc.
- Webster, E. C. (1982). The Employment Interview: A Social Judgment Process. Schomberg, Ontario: SIP Publications.
- Weekley, J. A., & Gier, J. A. (1987). Reliability and validity of the situational interview for a sales position. Journal of Applied Psychology, 72, 484-487.
- Wells, D. L., & Muchinsky, P. M. (1985). Performance antecedents of voluntary and involuntary managerial turnover. Journal of Applied Psychology, 70, 239-336.
- Wernimont, P. F., & Campbell, J. P. (1968). Signs, samples, and criteria. Journal of Applied Psychology, 52, 372-376.
- Wheeler, A. E., & Knoop, H. R. (1982). Self, teacher, and faculty assessments of student teaching performance. Journal of Educational Research, 75, 178-181.

- Wheeler, K. G. (1981). Sex differences in perceptions of desired rewards, availability of rewards, and abilities in relation to occupational selection. Journal of Occupational Psychology, 54, 141-148.
- Wheeler, K. G. (1984). Comparisons of self-efficacy and expectancy models of occupational preferences for college males and females. Journal of Occupational Psychology, 56, 73-78.
- Whitlock, G. H. (1963). Application of the psychophysical law to performance evaluation. Journal of Applied Psychology, 1, 15-23.
- Whitney, D. R. (1969). Predicting from expressed vocational choice: A review. Personnel and Guidance Journal, 48, 279-286.
- Wicklund, R. A. (1975). Objective self-awareness. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology (Vol. 8, pp. 233-275). New York: Academic Press.
- Wicklund, R. A., & Gollwitzer, P. M. (1983). A motivational factor in self-report validity. In J. Suls & A. G. Greenwald (Eds.), Psychological Perspectives of the Self (Vol. 2, pp. 67-92). Hillsdale, New Jersey: Lawrence Erlbaum.
- Wiggins, J. S. (1977). Personality and Prediction: Principles of Personality Assessment. Reading, Massachusetts: Addison-Wesley.
- Williams, W. E., & Seiler, D. A. (1977). Relationship between measures of effort and job performance. Journal of Applied Psychology, 57, 49-54.
- Working Group of the Scientific Affairs Board. (1986). Report of the working group on the use of the polygraph in criminal investigation and personnel screening. Bulletin of the British Psychological Society, 39, 81-94.
- Wright, O. R. (1969). Summary of research on the selection interview since 1964. Personnel Psychology, 22, 391-413.
- Zedeck, S. (1971). Problems with the use of moderator variables. Psychological Bulletin, 76, 295-310.

- Zedeck, S. (1986). A process analysis of the assessment centre method. In B. Staw and L. Cummings (Eds.), Research in Organisational Behaviour (Vol. 8, pp. 259-296). Greenwich, Connecticut: JAI Press.
- Zedeck, S., & Cascio, W. F. (1984). Psychological issues in personnel selection. Annual Review of Psychology, 35, 461-518.
- Zedeck, S., Tziner, A., & Middlestadt, S. E. (1983). Interviewer validity and reliability: An individual analysis approach. Personnel Psychology, 36, 355-370.

APPENDIX 1: LETTER SENT REQUESTING COOPERATION IN RESEARCH PROJECT.



Massey University

PSYCHOLOGY DEPARTMENT

PALMERSTON NORTH, NEW ZEALAND

TELEPHONES: 69-079, 69-089, 69-099

DATEX: NZ 30974, Mas Uni

In reply please quote:

6 June 1984

Dear

I am a PhD student at the above university conducting research on personnel selection. I am approaching different organisations in New Zealand requesting co-operation.

Briefly, my study is an evaluation of the contribution self-assessment can make to personnel selection procedures being used in an organisation. It involves attaching an easily administered self-assessment tool to a selection procedure presently used and then assessing selected individuals' job performance after they have been employed for a period of time. The usefulness of self-assessment, both in relation to the other selection methods, and as a predictor of job performance would then be evaluated.

Organisations involved will have access to an evaluation of their selection methods as well as use of the self-assessment tool should it prove of value. All data collected will be kept strictly confidential and no information to identify participating companies will be presented in the dissertation or elsewhere.

I would be very interested in conducting part of my study with . If this is possible could you please contact me at the above address. I will be available to come and see you at your convenience.

Yours sincerely

A handwritten signature in black ink, appearing to read 'Dave George', with a horizontal line extending to the right.

Dave George

**APPENDIX 2: SELF-ASSESSMENT REQUEST USED IN ELECTRONICS
MANUFACTURING COMPANY.**

Self-Assessment Form

for

Job Applicants

To the Job Applicant

This form is part of a university research project looking at personnel selection. _____ has agreed to ask people applying for jobs to fill in this form. Nothing you say here will be used for selection by the company.

This form gives you the chance to rate yourself on how well you think you will do this job. You may not have done work like this before but you will be able to give a 'best answer' using what you know about yourself, how you have worked in other jobs and what you know about the job.

When making a rating use the scale shown below:

POOR		AVERAGE		EXCELLENT
1	2	3	4	5

Read each item carefully. Decide on your rating by comparing yourself to the sort of people you think will be doing this type of work. Write your rating next to the question. Here is a completed example.

1. Timekeeping

How good are you at getting to work on time? 4

After filling in your name, today's date and the job applied for, go to the next page and answer all items. Please answer the questions carefully. If you are offered a job your ratings will be compared to how well you work.

Name _____ Date _____

Job Name _____

Time Keeping

How good are you at being on time for work in the morning as after tea and meal breaks? _____

Social Acceptability

How good are you at getting on with the people you are working with? _____

Social Interaction

How good are you at fitting in with a team of workers? _____

Work Area Tidiness

How good are you at keeping your work area tidy? _____

Care of Hand Tools

How good are you at caring for the hand tools you are using? _____

Work Attitude

How good are you at doing your job the way the company wants it done? _____

Care of Product

How good are you at taking care of the things you are making? _____

Work Speed

How good are you at getting work done on time? _____

Use of Tools

How good are you at using different hand tools? _____

Items 1 to 9 deal with some of the things involved in work at . Write in the spaces below which three of these items you think you would be good at doing and which three you would be not so good at doing.

GOOD AT DOING

NOT SO GOOD AT DOING

Overall Job Performance

How well do you think you can do the job you are applying for at ? _____

next few questions do not need a rating. Just tick the space next to each question for an answer.

Do you know what it is like to work for ? YES _____ NO _____

If your answer is YES how did you find out what it is like to work for ?
(Tick the answer which is true for you)

- (a) By working here before _____
- (b) By watching others work here _____
- (c) By being told by others (friends, relatives) what it is like _____
- (d) By seeing pictures, movies or reading about work here _____
- (e) Any other way you can think of? Write it in here: _____

Have you ever been asked to rate how well you could do a job by any employer other than ? YES _____ NO _____

If your answer is YES, was it when applying for that job? YES _____ NO _____

Or, while you were working for that firm? YES _____ NO _____

When you have completed all the questions hand this form to the personnel officer. Thank you for your help.

**APPENDIX 3: EMPLOYEE APPRAISAL FORMS USED IN ELECTRONICS
MANUFACTURING COMPANY.**

INDUCTION COURSE REVIEW						
REVIEW DATE: / /						
ITEM	ASSESSMENT					
	POOR					EXCELLENT
SOLDERING	1	2	3	4	5	N/A
COMPONENT RECOGNITION	1	2	3	4	5	
CARE OF PRODUCT	1	2	3	4	5	
COMPONENT INSERTION	1	2	3	4	5	
WORK SPEED	1	2	3	4	5	
USE OF TOOLS	1	2	3	4	5	
CARE OF HAND TOOLS	1	2	3	4	5	
WORK AREA TIDINESS	1	2	3	4	5	
TIMEKEEPING	1	2	3	4	5	
PERSONAL ATTITUDE	1	2	3	4	5	
SOCIAL HABITS	1	2	3	4	5	
TUTOR'S COMMENTS:						
/ /						
(SUPERVISORS SIGNATURE)			(EMPLOYEES SIGNATURE)		(DATE)	

4TH WEEK REVIEW

REVIEW DATE: 27 / 9 / 8

ITEM	ASSESSMENT				
	POOR	AVERAGE		EXCELLENT	OTHER
SOLDERING	1	2	(3)	4	5
MATERIAL RECOGNITION	1	2	(3)	4	5
CARE OF PRODUCT	1	2	(3)	4	5
COMPONENT INSERTION	1	2	(3)	4	5
WORK SPEED	1	2	(3)	4	5
USE OF TOOLS	1	2	(3)	4	5
CARE OF HAND TOOLS	1	2	(3)	4	5
WORK AREA TIDINESS	1	2	(3)	4	5
TIMEKEEPING	1	2	(3)	4	5
WORK ATTITUDE	1	(2)	3	4	5
SOCIAL INTERACTION	1	2	(3)	4	5
SOCIAL ACCEPTABILITY	1	2	3	4	(5)

GENERAL COMMENT:

*works well,
he's very quiet, but he
gets along with others quite
well.*

NOTE - IF ANY ITEM IS NOT APPLICABLE PLEASE ENTER N/A IN 'OTHER' COL.

F.A.

(SUPERVISORS SIGNATURE)

(EMPLOYEE'S SIGNATURE) (DATE)

7TH WEEK REVIEW

REVIEW DATE: 25 / 10 / 8

ITEM	ASSESSMENT				
	POOR	AVERAGE		EXCELLENT	OTHER
SOLDERING	1	2	3	4	5
MATERIAL RECOGNITION	1	2	3	4	5
CARE OF PRODUCT	1	2	3	4	5
COMPONENT INSERTION	1	2	3	4	5
WORK SPEED	1	2	3	4	5
USE OF TOOLS	1	2	3	4	5
CARE OF HAND TOOLS	1	2	3	4	5
WORK AREA TIDINESS	1	2	3	4	5
TIMEKEEPING	1	2	3	4	5
WORK ATTITUDE	1	2	3	4	5
SOCIAL INTERACTION	1	2	3	4	5
SOCIAL ACCEPTABILITY	1	2	3	4	5

GENERAL COMMENT:

*Still quiet,
but progressing well.*

SHOULD THIS EMPLOYEE BE RETAINED?

YES / ~~NO~~

(SUPERVISORS SIGNATURE)

DATE: *1 / 11 / 8*

SHOULD THIS EMPLOYEE RECEIVE A MERIT INCREMENT?

YES / NO

PLANT MANAGERS SIGNATURE:

DATE: / /

APPENDIX 4: INTERVIEWERS GUIDE AND GENERAL KNOWLEDGE TEST USED IN ELECTRONICS MANUFACTURING COMPANY.

JOB TITLE CANDIDATE'S NAME DATE

1. Attributes

2. Reservations

3. Typing and Skills Test

4. Reference Check

Score

1. Appearance	1	2	3	4	5	6	7	8	9	10
2. Manner/Personality	1	2	3	4	5	6	7	8	9	10
3. Communication Skills	1	2	3	4	5	6	7	8	9	10
4. Motivation	1	2	3	4	5	6	7	8	9	10
5. Experience	1	2	3	4	5	6	7	8	9	10

General Comments

SECTION 1 POSITION APPLIED FOR _____

LOCATION _____ DIVISION _____

NAME Mr/Ms _____
(Surname) (Christian Name/s)

ADDRESS _____

TELEPHONE _____

DATE OF BIRTH _____ MARITAL STATUS _____ HEALTH _____

NUMBER & AGE OF DEPENDANT CHILDREN _____

NAME & ADDRESS OF NEXT OF KIN _____

FULL TIME/PART TIME (Hours per day : _____)
(Total hours/week: _____)

DO YOU HAVE A CRIMINAL RECORD? _____ DO YOU SMOKE? _____

SECTION 2 EDUCATION

SCHOOL, COLLEGE, UNIVERSITY _____

AGE LEAVING _____ LEVEL ATTAINED _____

CERTIFICATES OR DEGREES _____

BEST SUBJECTS _____

RECREATIONS _____

OTHER ATTAINMENTS _____

SECTION 3 EXPERIENCE

PRESENT EMPLOYER POSITION FROM SALARY REASON FOR LEAVING

NOTICE REQUIRED BY PRESENT EMPLOYER _____

PREVIOUS EMPLOYERS	POSITION	STARTING DATE	FINISHING DATE	SALARY	REASON LEFT

DECLARATION: I declare that I am in good health, that all the above answers are to the best of my belief true and correct, and that a false declaration made on this form will result in instant dismissal.

SIGNATURE _____ DATE _____

INTERVIEWERS REMARKS :

COLOURBLIND TEST: _____

FACT SCORE: _____

EYE CHART TEST: _____

MEDIUM OF INTRODUCTION _____

PROPOSED STARTING DATE _____

PROBATION PERIOD/
REVIEW DATE : _____

SALARY OFFERED : _____

ELIGIBILITY TO WORK
IN NEW ZEALAND _____

EVIDENCE SEEN OF STATED SALARY LEVEL _____

SPECIAL CONDITIONS :

SIGNED :

1. Who is the Governor General of New Zealand?

ANSWER:

2. Who is the President of the Federation of Labour?

ANSWER:

3. Name the largest polynesian populated city in the world?

ANSWER:

4. Who is the leader of the New Zealand party?

ANSWER:

5. Who is the Company Personnel Officer for AWA?

ANSWER:

6. The Capital of New Zealand is Wellington. What city or town was the first Capital of New Zealand?

ANSWER:

7. In weight 2.2 pounds equals what metric measure?

ANSWER:

8. What voltage is the domestic power supply in New Zealand?

ANSWER:

9. What date is Waitangi Day?

ANSWER:

10. What is 's telephone number?

ANSWER:

APPENDIX 5: SUB-GROUP CORRELATIONS FROM ELECTRONIC ASSEMBLER SAMPLE.

Table A:

Sub-Group Correlations Between Self-Assessment Predictor (SAJOB), Organisational Selection Decision (GRADE) and Work Criteria (INDUCT, PERFORM).

<u>Moderator Grouping</u>	<u>Predictor - Criterion Relationship</u>		
	SAJOB/GRADE	SAJOB/INDUCT	SAJOB/PERFORM
Age			
< 22.6 years	0.19 (25)	0.17 (24)	-0.18 (18)
> 22.7 years	.	0.48 (21)	0.28 (18)
Social Desirability			
<22	0.12 (26)	0.46 (24)	-0.07 (19)
>23	.	0.04 (18)	0.16 (15)
Race			
Pacific Islander	-0.03 (14)	0.24 (13)	0.20 (19)
Maori	0.30 (19)	0.32 (16)	0.13 (15)
Caucasian	0.15 (16)	0.52 (14)	0.40 (10)
Job Knowledge			
Enactive	.	0.99 (3)	0.37 (6)
Vicarious	.	.	.
Being told	0.30 (19)	0.32 (19)	-0.04 (13)
Books, pictures, etc.	.	-0.75 (5)	0.06 (4)
No source	0.09 (19)	0.40 (17)	0.13 (12)
SA experience			
No	0.15 (38)	0.29 (32)	-0.03 (27)
Yes	0.10 (13)	0.35 (13)	0.05 (9)
Total Sample	0.14 (51)	0.31 (45) #	0.18 (43)

Notes

1. Values of n for each sub-group shown in parentheses.
2. # = correlation significant at $p < .02$. No significant differences between sub-groups was detected.
3. Where a correlation cannot be calculated due to a lack of group numbers or no variability in one or other of the variables '.' is shown.
4. In the case of two sub-groups a test of significant differences between correlations was used (Edwards, 1984). Where more than two coefficients were involved a test for the homogeneity of several correlations was used (Edwards, 1984).

Table B:

Sub-Group Correlations Between Average of Self-Assessment Ratings of Work Aspects (SELFAV), Organisational Selection Decision (GRADE) and Work Criteria (INDUCT, PERFORM).

<u>Moderator Grouping</u>	<u>Predictor - Criterion Relationship</u>		
	SELFAV/GRADE	SELFAV/INDUCT	SELFAV/PERFORM
Age			
< 22.6 years	0.21 (31)	0.12 (30)	-0.21 (21)
> 22.7 years	.	-0.16 (25)	0.04 (22)
Social Desirability			
<22	0.20 (29)	0.26 (27)	-0.33 (20)
>23	0.03 (27)	-0.12 (23)	-0.07 (19)
Race			
Pacific Islander	0.05 (18)	-0.23 (17)	-0.12 (14)
Maori	0.28 (22)	0.35 (19)	0.27 (16)
Caucasian	-0.16 (19)	0.15 (17)	-0.35 (12)
Job Knowledge			
Enactive	0.08 (7)	-0.12 (3)	-0.25 (6)
Vicarious	.	.	.
Being told	0.24 (21)	0.15 (21)	-0.13 (14)
Books, pictures, etc.	.	-0.59 (7)	-0.57 (12)
No source	0.03 (24)	0.00 (22)	0.07 (16)
SA experience			
No	0.06 (44)	-0.12 (38)	-0.23 (30)
Yes	0.20 (17)	0.10 (17)	0.15 (13)
Total Sample	0.10 (61)	-0.03 (55)	-0.07 (43)

Notes

1. Values of n for each sub-group shown in parentheses.
2. # = correlation significant at $p < .02$. No significant differences between sub-groups was detected.
3. Where a correlation cannot be calculated due to a lack of group numbers or no variability in one or other of the variables '.' is shown.
4. In the case of two sub-groups a test of significant differences between correlations was used (Edwards, 1984). Where more than two coefficients were involved a test for the homogeneity of several correlations was used (Edwards, 1984).

APPENDIX 6: SELF-ASSESSMENT REQUEST USED IN THE FOOD PROCESSING COMPANY (STUDY A).

NUMBER DATE

CT CREDIT ACCOUNT NO:

AME

T NAME

OF BIRTH SEX FEMALE/MALE (Delete one)

PHONE NUMBERS HOME/CONTACT WORK

RESS 1.

2.

YOU WORKED FOR ' BEFORE? YES NO IF YES - WHEN _____ AND IN WHAT JOB _____

WH SHIFT DO YOU PREFER? - NIGHT(10.00pm/6.00pm) MORNING(6.00am/2.00pm) AFTERNOON(2.00/10.00pm)

YOU ABLE TO WORK SEVEN DAYS A WEEK? YES NO

YOU SUFFER ANY DISABILITY WE SHOULD KNOW ABOUT AND IF SO WHAT? _____

BEEN CONVICTED FOR ANY OFFENCE? _____ DO YOU HAVE PROCEEDINGS PENDING? _____

OR HIGH SCHOOL EDUCATION - HOW MANY YEARS? _____ STANDARD REACHED? _____

EMPLOYMENT

EMPLOYER'S NAME & ADDRESS	POSITION & DUTIES	DATES START & FINISH	REASON FOR LEAVING

I declare that to the best of my knowledge the information on this form is correct and acknowledge that this information is provided on the basis of my employment at _____ Ltd. I authorise the deduction of Union Fees from my wages by the appropriate union. Signature _____
 This information will be retained for the period of the current season. Should you still be interested at that time please contact us again.

is very interested in how well you think you can do this work. This part of the form gives you the chance to rate yourself on parts of the job as well as some personal abilities the work may require.

When answering those questions which ask for a rating, compare yourself with the sort of people you think would be doing this type of work. Even though you may not have done this job before, you will be able to give good answers using what you know about yourself, what you have learned about this job and past experience in other jobs. The scale to be used for your ratings is shown below.

1. An excellent standard of performance.
2. A high standard of performance.
3. A reasonable standard of performance.
4. A not quite acceptable standard of performance.
5. A poor standard of performance.

Read the question carefully then choose your rating from the scale by comparing yourself to the sort of people you think would be doing this type of work. Here is a completed example.

3. How good are you at getting to work on time?

4

Now go to the opposite page and answer all the questions after filling in your name and today's date. Please answer the questions carefully.

There are many different jobs in the canneries. Listed below are seven types of work done by workers during the coming season. We are interested in how well you think you would do each type of work if it was to be your job for the season. Read each description carefully then give yourself a rating showing how well you would do that type of work.

1. Inspection

Inspects fruit or vegetables as they pass along a belt. Fruit and vegetables that are marked or damaged are removed along with any foreign matter such as sticks or leaves. Inspection may also involve cans, where damaged or poorly labelled cans are removed.

RATING _____

2. Trimming

Fruit or vegetables are taken from a moving belt and trimmed to remove marked or damaged parts.

RATING _____

3. Beltwork

This involves working on the production line doing a number of things by hand. Some examples are: painting carton flaps with glue, filling cans, cartons or packets with fruit or vegetables, labelling cans, and putting cans into crates.

RATING _____

4. Machine Feeding

Feeding fruit and vegetables into processing machines. An example is placing apples or pears on the feedcups of peeling machines.

RATING _____

5. Cleaning

Carrying out general cleaning, tidying and washing down duties.

RATING _____

6. Machine Operating

Operating machines to process fruit or vegetables. Some machines involved are those which close cartons, seam cans, bag frozen vegetables, put cans on pallets and control the flow of juice, fruit and vegetables.

RATING _____

7. Heavy Line Work

This involves such work as carrying, stacking and unloading bags and cartons, sewing large bags, emptying bins or cases, putting together cardboard bins etc.

RATING _____

8. Which two of the above types of work do you think you would be good at and which two types would you be poor at doing. Write them in the spaces provided below.

GOOD AT DOING

POOR AT DOING

9. Give yourself a rating showing how good you are at being at work on all working days.

RATING _____

10. Overall Job Performance

How well do you think you will do the job you are given in the canneries?

RATING _____

The next few questions do not require a rating. Just put a tick in the space next to your answer.

11. Do you know what it is like to work in the canneries?

YES _____ NO _____

If your answer is YES, how did you find out what it is like to work in the canneries? (Tick each of the answers which are true for you).

- By working in the canneries myself _____
- By watching others working in the canneries _____
- By being told by others what it is like to work in the canneries _____
- By reading, seeing pictures, movies, etc of work in the canneries _____
- If you learnt about working in the canneries from somewhere or somebody not mentioned above, please write that in here _____

12. Do you know what sort of job you are going to be doing in the canneries

YES _____ NO _____

If YES name the job. _____

13. Have you ever been asked to rate your performance by any other employer?

YES _____ NO _____

- If your answer is YES has it been while applying for another job? YES _____ NO _____
- Has it been while working for that employer? YES _____ NO _____

When you have completed all the questions hand this form to the Personnel Officer. Thank you for your application.

APPENDIX 7: SELF-ASSESSMENT REQUEST USED IN THE FOOD PROCESSING COMPANY (STUDY B).

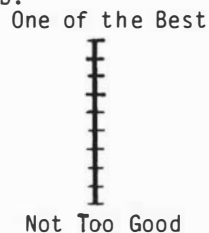
SEASONAL APPLICATION FORM

NAME _____

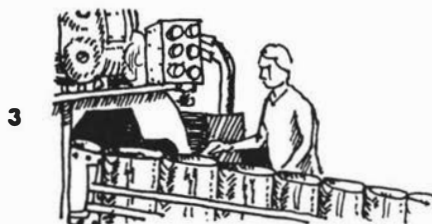
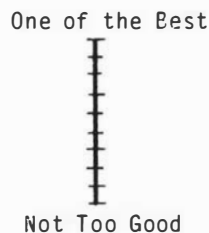
want to know how you think you will work during the season compared to other seasonal workers. There are a number of different seasonal jobs. They are named and described below. Show what sort of worker you would be in the jobs shown below by putting a cross on the scale for each job.



Driving
Driving forklifts, trucks or harvester mills.



Belt Work
Working on belts, trimming and inspecting fruit and vegetables, working with cartons and feeding on cans.



Machine Operating
Controlling machines to process fruit or vegetables. E.g. closing cans, putting cans on pallets, controlling the flow of fruit, juice, etc.



Field Work
Planting and harvesting tomatoes



General Line Work
Carrying, stacking and filling bags, cartons, and bins. Emptying bins. Putting large bins together. Cleaning up.



6. Out of Driving, Belt Work, Machine Operating, Field Work and General Line Work which ONE or TWO would you do best and which ONE or TWO would you do poorly?

BEST AT _____ POOR AT _____

PLEASE TURN THE PAGE 

7. If you know what it is like to work for _____ during the season how did you find out about the work? (Tick the answer which is right for you)

- By working for myself _____
- By watching others work at _____
- By being told by others about work at _____
- Any other way? (write here) _____

Please print the answers to the following questions.

- 8. Can you work 7 days a week if required? _____
- 9. Do you suffer from any disability or allergy? _____
- 10. Have you been convicted of any offence? _____
- 11. Have you a court appearance coming up? _____
- 12. How many years did you attend High School? _____
- 13. What was the highest qualification you gained? _____

14. Who were your last two employers?

	Last employer	Second to last employer
Name and Address	_____	_____
Your Position	_____	_____
Length of Service	_____	_____
Reason for Leaving	_____	_____

15. IF YOU HAD A CHOICE of jobs among Driving, Beltwork, Machine Operating, Field Work or General Line Work, which job would be:

- your 1st Choice _____
- your 2nd Choice _____
- your 3rd Choice _____
- your 4th Choice _____

I certify to the best of my knowledge that the information on this form is correct and acknowledge that this information will form the basis of my employment at _____ I authorize the deduction of union fees from my wages for the appropriate union.

Signature _____

This form will be kept for the current season. Should you be interested after that time please contact us again.



APPENDIX 8: SELF-ASSESSMENT REQUEST USED IN THE POLYTECHNIC NURSING COURSE (STUDY A).

Dear Applicant,

In order to help us in considering your application, would you please return this form to us, having filled in your own assessment of yourself on the attributes listed.

We also enclose a reference form, together with a stamped, addressed envelope, be given to the referee named in your returned application (the principal of your school or your employer). Please ensure that the referee receives this promptly so that your application may be processed as quickly as possible.

- 3) Personal qualities (give your opinion of yourself for each subsection)
- 4) Write how you think you relate to people in these different groups.
- 5) Again, write your assessment of your attitudes to work.
- 6) Add any additional comments you think may assist us in assessing your suitability for a nursing career.

Please sign and date this form.

NAME: _____

3. Personal qualities:

Integrity: _____

Honesty: _____

Acceptance of responsibility: _____

Health: _____

Personal appearance: _____

Special abilities: _____

4. Comment on the applicant's relationships with other people:

with people of their own age group _____

with people in authority: _____

5. Attitudes to assigned tasks or work:

ability to accept correction and direction: _____

co-operation with others: _____

perseverance: _____

concern for the needs of others: _____

6. Do you think he/she has suitable personal qualities for a nursing career?

7. Do you think he/she has the ability to cope with the academic demands of a 3-year full-time course?

8. Additional comments: _____

Signature: _____

Date: _____

APPENDIX 9: SELF-ASSESSMENT REQUEST USED IN THE POLYTECHNIC NURSING COURSE (STUDY B).

PERSONAL QUALITIES

Please place an X in the space on the scale which best indicates your assessment of the applicant in relation to each of the following qualities:

- | | | |
|------------------------|---|-----------------------|
| Honest | : _____ : _____ : _____ : _____ : _____ : | Dishonest |
| Mature | : _____ : _____ : _____ : _____ : _____ : | Immature |
| Reliable | : _____ : _____ : _____ : _____ : _____ : | Unreliable |
| Well groomed | : _____ : _____ : _____ : _____ : _____ : | Untidy |
| Tolerant | : _____ : _____ : _____ : _____ : _____ : | Intolerant |
| Accepts responsibility | : _____ : _____ : _____ : _____ : _____ : | Avoids responsibility |

Please comment on your assessment of this applicant's personal qualities: _____

INTERPERSONAL RELATIONSHIPS

Please comment on the following:

Relationship with peers: _____

Relationship with people in authority: _____

Consideration of others: _____

ATTITUDES TO WORK/STUDY

Please comment on the following:

Perseverance: _____

Co-operation with others: _____

Application to work/study: _____

Acceptance of correction: _____

Initiative: _____

GENERAL

Has the applicant's health ever affected his/her performance at work/school? YES NO

If YES, comment: _____

Is attendance pattern acceptable? YES NO

If NO, comment: _____

Does applicant have special abilities or disabilities? _____

Do you consider the applicant able to undertake a full time 3 year course of Intensive study? YES NO

Give reason(s): _____

Signed: _____ Position: _____ Date: _____

The following information is provided to assist you in completing the enclosed form. It is based on both local and national information and statistics.

Possession of the personal qualities listed on the left hand side of the enclosed form is a primary requisite for success in nursing. Good health is another essential, since both the course itself and a nursing career are demanding. A nurse must also be able to work well with others. Interpersonal relationships are the foundation for all nursing work and although specific skills are taught within the course, it is unlikely that the required standard of therapeutic skill can be reached if the student does not come with well developed social skills and indications that they do care for others.

The three year Comprehensive Nursing course is demanding in terms of work and study. In our first year course, students must cope with a wide variety of theoretical studies at once as well as learning the many practical skills and attaining a certain level of mastery in these. Examples of the theoretical subjects which are all taken concurrently include Anatomy and Physiology; Chemistry; Microbiology; Calculation Skills; Sociology; Psychology; Anthropology; Human Growth and Development; Communications; as well as the major subject, Nursing Theory.

The following statistics from the Department of Health Returns for 1983 give the NATIONAL picture of the highest educational attainment for entrants into Comprehensive Nursing courses in New Zealand.

Advanced University Degree	0.3%
University Degree - Bachelors	2.5%
University or Teachers Diploma	0.5%
University Bursary/Scholarship	11.7%
Higher School Certificate	13.4%
University Entrance (i.e. 4 passes)	55.7%
6th Form Certificate	6.5%
Minimum Educational Qualifications as defined in 8a of Nurses' Regulations 1979	1.1%
Enrolment as a Nurse	2.7%
Mature Age Entry as defined in 8b of Nurses' Regulations 1979	5.7%

Essentially it is becoming very difficult for a school leaver to gain entry to a Comprehensive Nursing Course without having at least 4 passes in U.E. subjects and the trend seems to be towards taking students from 7th form or after a year away from school rather than taking them straight from 6th form. Exceptions to this would generally only be for students who had already demonstrated personal attributes which are valued in nursing, maturity and above average academic ability.

APPENDIX 10: SELF-ASSESSMENT REQUEST USED IN THE GOVERNMENT DEPARTMENT (STUDY A).

Self-Assessment Form

for

Job Applicants

NAME _____

DATE _____

The Department _____ is very interested in how well you think you can do the sort of work a graduate recruit would be doing. This form gives you the opportunity to rate yourself on different aspects of the position as well as some personal qualities the work may require. Although you may have had little or no experience in this type of position you will be able to give a 'best answer' based on knowledge of yourself, similar past experience and what you have learnt about the position applied for.

Read each item carefully. Give yourself a rating that indicates how well you feel you would perform on each item compared to the sort of person you think would be doing this type of work.

Please give careful consideration to your ratings. If you are offered this job the ratings you give now will be compared to your work performance in the future.

1. ORGANISATION OF WORK

- Outstanding organising ability
 Very effective
 Effective
 Needs Development
- ANY COMMENT?

2. QUALITY OF WORK (Consider your standard of work, presentation, accuracy and workmanship)

- Very high standard
 High standard
 Acceptable standard
 Need to improve
- ANY COMMENT?

3. OUTPUT

- Achieve an outstanding amount of work
 Achieve a high output
 Achieve required output
 Need to improve (Give reasons)
- ANY COMMENT?

4. STAFF MANAGEMENT (Consider whether you get the best out of people and how this is achieved)

- Very effective
 Effective
 Needs improvement
 Not applicable
- ANY COMMENT?

5. COMMUNICATION (Consider your ability to express yourself clearly and concisely - orally and/or in writing)

- Highly effective
 Effective
 Need to improve
- ANY COMMENT?

6. JUDGEMENT (Consider whether you are able to assess situations and develop appropriate solutions)

- Very sound
 Sound
 Require improvement
- ANY COMMENT?

7. VERSATILITY (Consider your ability to cope with a variety of work)

- Very successful
 Cope adequately
 Have difficulty
 Not tested
- ANY COMMENT?

8. INITIATIVE (Consider whether you see when something needs doing and do it, and whether you demonstrate innovative)

- Act without prompting
 Generally require no prompting
 Often have to be told what to do
- ANY COMMENT?
-

3. DEPENDABILITY

Very dependable
ANY COMMENT?

Can be depended on

Need extra supervision

4. STAFF RELATIONS (Consider how you work with staff of all levels and the effect this has on job performance)

Very good
ANY COMMENT?

Good

Need improvement

1. PUBLIC RELATIONS (Consider how you deal with the public, people other than those you work with)

Very good
ANY COMMENT?

Good

Need improvement

Not test

2. Items 1 to 11 deal with some aspects of work. Write in the spaces provided below up to four of these aspects you feel most competent in and up to four aspects you feel least competent in.

MOST COMPETENT

LEAST COMPETENT

3. ASSESSMENT DAY RATING

If you had to rate yourself on how well you would do the job, based on your performance over the assessment day, what rating would you have given yourself? (Circle one).

POOR C- C C+ B- B A- A A+ EXCEPTIONAL

4. OVERALL JOB PERFORMANCE

How effective do you think you will be in fulfilling the major duties, responsibilities and objectives of the job. (Circle one).

- 1 A standard of performance rarely achieved by others
- 2 A standard of performance substantially exceeding job requirements
- 3 A standard of performance fully meeting the requirements of this job
- 4 A standard of performance not entirely meeting all requirements of the job
- 5 A standard of performance which is of limited effectiveness

The final items do not require a rating. Just respond to the questions as asked.

15. Have you ever been asked to assess your performance by any other employer? Yes/No

If your answer is YES, has this request for self assessment been:

a) While working for that employer, such as a performance appraisal? Yes/No

b) When applying for a job with that employer? Yes/No

When you have completed all the items and filled in your name and the date on the front cover, please hand the form to the personnel officer. Thank you for your co-operation.

APPENDIX 11: DESCRIPTION OF GOVERNMENT DEPARTMENT ASSESSMENT CENTRE COMPONENTS.

ASSESSMENT CENTRE CONTENT

Introductory Session

The introductory session outlined the objectives of the Assessment centre and provided background information on the organisation.

Question in the House Work Sample

This exercise required candidates to identify the main elements of a question asked in the House and provide an outline of a possible answer. They were provided with a copy of the newspaper article on which the question is based, a comment on events in Korea from the "Deer Farmer", a background report on export of deer velvet and deer antler products and a handout "Questions to Ministers" giving the official view on questions for ministers.

The "question" was:

Has the minister seen the article in the Evening Post of Wednesday 12 September in which it is alleged that New Zealand exporters of deer velvet are adulterating their products to increase their returns? Is the minister concerned at the report and what is his government doing to prevent these practices and so protect this country's image as a supplier of quality products?

Import Licence Problem

Candidates had to make a decision on whether or not a company proposing to expand its product range should be allowed to import the components this expansion requires. For this candidates were provided with an outline of the Global Tender Scheme, a description of the proposal and a recent press release describing government procedure and policy. Candidates had 30 minutes. There was no "right" answer but a number of good reasons for a decision either way.

Import Licence Meeting

This exercise uses information from the import licence problem. It involved a group discussion where candidates were asked to explain and defend their decisions and reasons they committed to paper in response to the import licence question. Advisory Officers are often required to attend meetings both within and outside the department. At these meetings they are also required to represent a particular point of view based on policy, department decisions, interpretation of various political requirements, etc. The aim of this group discussion was to simulate such a meeting.

Press Release Work Sample

Candidates were assigned the role of an advisory officer in the Economics Directorate. They had been asked by their Executive Officer to prepare a Press Release on the Consumers Price Index (CPI) for the September Quarter. They were provided with a Consumer Price Index - September Quarter, a handout titled "Ministerial replies and speeches on prices: Themes for inclusion" and an example of a press release "Industry Assistance Reform".

Report Outline Work Sample

Candidates "played" the role of a newly employed Assistant Advisory Officer working in the Industrial Development Division of the organisation who has been assigned responsibility for monitoring the New Zealand Aluminum Smelter at Tiwai Point.

The smelter company had requested special consideration from the government in its energy supply agreements. The minister asked for a report. This report was to be written in full by the Executive Officer. To assist, the "desk officer" was requested to outline the factors they feel should be considered in preparing an initial report for discussion within the department. Candidates were given a profile of the company produced by NZ Aluminum Smelters, a transcript of a radio programme containing the Prime Minister's reaction, a press release from the office of the Ministry of Energy and a statement from the opposition spokes person on the smelter situation as well as a series of newspaper articles relating to the topic.

Ministerial Letter Work Sample

The candidate was assigned the role of a desk officer who has been asked by their Executive Officer to draft a reply to a letter received by the a Minister. The "desk officer" was instructed to deal with the issues raised so there is no call for further correspondence, to write the reply in the context of present government policy, and to satisfy the writer. They were provided with the letter, a copy of a newspaper cutting "CER solution in view" and Customs Tariff Schedules 3 & 5.

Ministerial Speech Work Sample

Candidates were instructed to watch a video taped presentation of a ministerial speech and then provide a brief outline of the main points made by the speaker. The presentation lasted approximately 25 minutes, followed by 15 minutes allowed for writing the summary. This exercise simulated a situation often experienced by desk officers where they are required to attend a talk, conference, meeting, etc and report clearly and concisely important aspects of the event.

Leaderless Group Discussion

For the purposes of the assessment centre this exercise was designed to observe how candidates performed in a group situation where there is no assigned leader but a requirement for co-operation in solving a task. The Leaderless Group Discussion (LGD) simulated a number of situations staff at desk officer level and above face regularly in their work. One example is an ad hoc committee formed to study particular problems, advise on a policy or event, design a training course, etc.

This particular exercise involved assigning the eight candidates the role of a committee of General Managers. The "committee" had been called together to select a General Manager for a new plant. The members were instructed by the coordinator to choose the correct candidate for the executive position based upon the data they are given. That data consists of candidate summary sheets, briefing sheets and individual data sheets. No one person alone had enough data to provide a solution to the problem.

Panel Interview

The panel interview was conducted by three departmental executives. It followed a standardized format. The information requested of candidates was to be based on job requirements as identified by job analysis. In addition, questions that desk officers often have to deal with in client interview situations, are asked.

Critical Thinking Appraisal

The Watson-Glaser Critical Thinking Appraisal is a series of test exercises which require the application of some of the important abilities involved in critical thinking. The exercises include problems, statements, arguments and interpretations of data similar to those encountered in daily life.

Meeting Desk Current Desk Officers

A group of approximately eight current desk officers were invited to lunch with the candidates. This was to give the candidates an opportunity to meet individuals working in the department, to ask questions about the work and to be shown around the work areas where the desk officers involved are stationed.

APPENDIX 12: SELF-ASSESSMENT REQUEST USED IN THE GOVERNMENT DEPARTMENT (STUDY B).

POST-ASSESSMENT CENTRE QUESTIONNAIRE FOR GRADUATE RECRUITS

The Department of _____ is interested in how well you think you can do the sort of work a graduate recruit does. This form gives you the opportunity to rate yourself on different tasks that are often part of their job now you have some experience of work in the department.

Read each item carefully. Give yourself a rating that indicates how you feel you would perform on each task. Please give careful consideration to your ratings. If you are offered a job, the ratings you give will be compared to your work performance in the future.

ASSESSMENT CENTRE RATING

If you had to rate yourself on how well you would do the job based on your performance over the assessment centre what rating would you give yourself? (Circle your rating)

Very Poor |-----|-----|-----|-----|-----|-----|-----|-----|-----| Exceptional
Performance 1 2 3 4 5 6 7 8 9 10 Performance

Any Comment? _____

Each of the following items describe a task that is often part of a graduate recruit's work. Rate the items as outlined above. In the space under each item please quote the experience on which you base your ratings.

- How would you perform if asked to answer a parliamentary question asked of a Government Minister by an Opposition MP. You would have access to relevant background information. Your answer would need to be in the form of a brief outline that accurately answers such a question in a positive manner. (Circle your rating)

Very Poor |-----|-----|-----|-----|-----|-----|-----|-----|-----| Exceptional
Performance 1 2 3 4 5 6 7 8 9 10 Performance

Experience on which you base this rating _____

- How would you perform if asked to provide a decision and the reasons behind that decision on an application by an organization to Trade and Industry for an import licence. There would probably be many good reasons for a decision either way. You would have access to relevant background material. (Circle your rating)

Very Poor |-----|-----|-----|-----|-----|-----|-----|-----|-----| Exceptional
Performance 1 2 3 4 5 6 7 8 9 10 Performance

Experience on which you base this rating _____

4. How would you perform if asked to produce a press release on a Quarterly Consumer Price Index (CPI). The release would need to be a professional interpretation of the CPI to capture media attention while reflecting positively, current political themes. (Circle your rating)

Very Poor |-----|-----|-----|-----|-----|-----|-----|-----|-----| Exceptional
Performance 1 2 3 4 5 6 7 8 9 10 Performance

Experience _____

5. How would you perform if asked to produce an outline for an internal report on a specific industry from background information such as files, company information sheets, newspaper and radio reports and Trade and Industry memoranda. The report would need to accurately reflect the current situation in relation to the specific problem which prompted the report. (Circle your rating)

Very Poor |-----|-----|-----|-----|-----|-----|-----|-----|-----| Exceptional
Performance 1 2 3 4 5 6 7 8 9 10 Performance

Experience on which you base this rating _____

6. How would you perform if asked to reply to a letter written to a Government Minister by a private individual or organization. The reply would need to deal with the issues in the letter so as to avoid further correspondence while reflecting government policy and endeavouring to satisfy the writer. (Circle your rating)

Very Poor |-----|-----|-----|-----|-----|-----|-----|-----|-----| Exceptional
Performance 1 2 3 4 5 6 7 8 9 10 Performance

Experience on which you base this rating _____

7. How would you perform if asked to construct a summary of a political speech for departmental officials who are unable to be present to hear the speech themselves. The summary would need to be an accurate representation of the speech. (Circle your rating)

Very Poor |-----|-----|-----|-----|-----|-----|-----|-----|-----| Exceptional
Performance 1 2 3 4 5 6 7 8 9 10 Performance

Experience on which you base this rating _____

7. How would you perform if asked to represent and support a decision you had made on an import licence application in a meeting. There would probably be many good reasons for a decision either way on the import licence application. There may be both opposition and support for your decision in the meeting. (Circle your rating)

Very Poor |-----|-----|-----|-----|-----|-----|-----|-----| Exceptional
Performance 1 2 3 4 5 6 7 8 9 10 Performance

Experience on which you base this rating _____

Items 1 to 7 deal with some of the tasks you may encounter when working in Trade and Industry. Write in the spaces provided which two or three tasks you feel most competent at and which two or three you feel least competent at.

MOST COMPETENT

LEAST COMPETENT

8. OVERALL JOB PERFORMANCE

How effective do you think you will be in fulfilling the major duties, responsibilities and objectives of the job. (Circle your rating)

Very Poor |-----|-----|-----|-----|-----|-----|-----|-----| Exceptional
Performance 1 2 3 4 5 6 7 8 9 10 Performance

Experience on which you base this rating _____

Do you have any comments on the Assessment Centre?

NAME _____ DATE _____

When you have completed all the items and filled in your name and today's date above, please hand the questionnaire to the Assessment Centre Co-ordinator. Thank you for your co-operation.

APPENDIX 13: SUB-GROUP CORRELATIONS FROM THE DESK OFFICER SAMPLE (GOVERNMENT DEPARTMENT STUDY B).

Table A:

Sub-Group Correlations Between Self-Assessment Predictors and Job Performance Ratings Using Self and Superior Sources.

<u>Moderator Grouping</u>	<u>Predictor - Criterion Relationships</u>			
	<u>Presa/Postsa</u>	<u>Postsa/Pasa</u>	<u>Postsa/Perform</u>	<u>Postsa/Promote</u>
Age				
<22.3 yrs	0.52 (43)	0.09 (25)	0.01 (25)	-0.05 (25)
>22.3 yrs	0.61 (42)	0.28 (21)	-0.30 (21)	-0.05 (22)
Critical Thinking				
<76	0.55 (43)	0.10 (21)	-0.01 (21)	0.11 (22)
>77	0.59 (42)	0.33 (25)	-0.44 (24)	-0.17 (25)
Social Desirability				
<17	0.63 (43)	0.43 (23)	-0.33 (23)	0.04 (24)
>18	0.54 (42)	0.04 (23)	0.09 (23)	-0.07 (23)
Gender				
Female	0.55 (46)	0.06 (29)	-0.02 (29)	-0.16 (29)
Male	0.61 (39)	0.54 (17)	-0.43 (17)	0.15 (18)
Work experience				
No full time job	0.68 (11)	-0.46 (5)	-0.05 (5)	-0.41 (5)
Full time job	0.56 (74)	0.19 (41)	-0.10 (41)	-0.02 (42)
Post-Graduate Degree				
No graduate degree	0.56 (47)	0.09 (25)	-0.20 (25)	-0.18 (25)
Graduate degree	0.60 (38)	0.37 (21)	-0.39 (21)	0.11 (22)
SA based on work experience				
No	0.53 (53)	0.11 (31)	-0.27 (31)	-0.03 (31)
Yes	0.54 (32)	0.33 (15)	0.24 (15)	-0.06 (16)
SA experience				
No	0.56 (70)	0.25 (36)	-0.03 (36)	-0.07 (37)
Yes	0.61 (15)	0.04 (10)	-0.46 (10)	-0.04 (10)
Total Sample	0.57 (85)	0.19 (46)	-0.14 (46)	-0.05 (47)

Notes

1. Values of n for each sub-group shown in parentheses.
2. * = Difference between sub-group correlations significant at $p < .05$
3. **Presa** = Pre-assessment centre self-assessment; **Postsa** = Post-assessment centre self-assessment; **Pasa** = Self-assessed performance appraisal; **Perform** = Superior performance appraisal; **Promote** = Promotion (based on salary steps awarded).

Table B:

Sub-Group Correlations Between Concurrent Self-Assessment and Job Performance Ratings Using Self and Superior Sources.

Moderator Grouping	Postsa/Offer	Pasa/Perform	Pasa/Promote
Age			
<22.3 yrs	-0.04 (43)	0.16 (25)	0.51 (25)
>22.3 yrs	0.15 (42)	0.23 (21)	0.26 (21)
Critical Thinking			
<76	0.04 (43)	0.45 (21)	0.60 (21)
>77	-0.06 (42)	-0.08 (25)	0.10 (25)
Social Desirability			
<17	-0.08 (43)	0.07 (23)	0.38 (23)
>18	0.05 (42)	0.27 (23)	0.25 (23)
Gender			
Female	0.02 (46)	0.39 (29)	0.41 (29)
Male	0.16 (39)	-0.10 (17)	0.29 (17)
Work experience			
No full time job	-0.10 (11)	0.69 (5)	0.28 (5)
Full time job	0.09 (74)	0.19 (41)	0.38 (41)
Post-Graduate Degree			
No graduate degree	-0.10 (47) *	0.22 (25)	0.26 (25)
Graduate degree	0.34 (38)	0.13 (21)	0.53 (21)
SA experience			
No	0.10 (70)	0.23 (36)	0.50 (36)
Yes	0.05 (15)	-0.07 (10)	-0.15 (10)
Total Sample	0.07 (85)	0.19 (46)	0.36 (46)

Notes

1. Values of n for each sub-group shown in parentheses.
2. * = Difference between sub-group correlations significant at $p < .05$
3. **Offer** = Job offer made following assessment centre; **Postsa** = Post-assessment centre self-assessment; **Pasa** = Self-assessed performance appraisal; **Perform** = Superior performance appraisal; **Promote** = Promotion (based on salary steps awarded)