AN EVALUATION OF SELF ASSESSMENT ON PERSONALITY TESTS FOR PERSONNEL SELECTION.

A thesis presented in partial fulfilment of the requirements for the degree of Master of Arts in Psychology at Massey University.

NICOLA. M. MORGAN.

1992

ABSTRACT

Personality tests are used commonly throughout the world as a tool for personnel selection. However, such use of these tests has caused much controversy among psychologists, with a number of researchers finding they have poor validity in predicting job performance. The present study proposed to cast further doubt on the use of personality tests for personnel selection by showing that people are able to predict their personalities as measured by personality tests, and can do so without responding in a socially desirable way.

Three experiments were performed. In experiment one the subjects did the Sixteen Personality Factors Questionnaire (16PF) and did their self assessments on the 16PF profile form. The subjects in experiment two followed the same procedure using the California Psychological Inventory (CPI). The subjects in the third experiment were asked to describe an ideal person via the use of the profile forms for the 16PF and CPI. These results were used to see whether the subjects in experiment one and two had responded in a socially desirable way when doing their self assessments. It was found that people are not very accurate at evaluating their personalities as measured by personality tests. Social desirability was not a major problem. For the 16PF five significant correlations were found between subjects' self assessment and personality test scores ($p < 0.01$) and nine factors were found to be responded to in a socially desirable way. For the CPI this was six and five respectively.

No moderators were found. It is concluded that although subjects were not very accurate at assessing their personalities as measured by personality tests, doubt is cast over their use as a selection tool, because the question of which is more accurate, the personality test or the person's self assessment has not been answered.

ACKNOWLEDGEMENTS.

Thank you to my supervisor Dr Mike Smith for encouragement and advice which was needed and heeded throughout this study. He has guided me and taught me much in the time that I have known him.

My thanks to the staff in the Department of Psychology; Dr John Spicer for advice on statistical matters, Dr Ross St George and Dr John Podd for their help with administration matters and Kate Mirfin, my mentor, for being there when I needed her. Thank you too to the students who volunteered as subjects. I could not have done it without you all.

Finally my thanks for the emotional support and encouragement I got from my friends, especially Paddy Mansell, my best friend, and my mother. Thanks too to Martin, a close friend who offered a number of distractions but was always there when I needed him over the last year.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

CHAPTER ONE

INTRODUCTION

1.01 PERSONNEL SELECTION

The heart of the selection process is in prediction. Although there are a number of selection procedures currently used it has been argued that the most valid selection methods are those which have a strong relationship to the content of the job (Ascher and Sciarrino, 1974; Smith and George, 1992).

Research by Dakin and Armstrong (1989) found however that selection methods recommended in research evidence did not correlate with consultants' beliefs of valid selection methods or with their practices.

A paper presented by Pearlman (1985) outlined a number of popular selection methods and their validities (see table one).

There is limited research on personnel selection practices in New Zealand (Inkson 1987). The research that has been conducted indicates that personnel selection generally involves unstructured interviews, application forms and references even though the validity of such instruments has been found to be poor.

Table One

Average Validity of Predictors in the Early 1980s.

Predictor Type	Validity
work sample/ performance/ skill test	.54
ability test	.53
assessment centre/ job simulation	.50
peer evaluation	.49
behavioural consistency experience rating	.49
job knowledge test	.48
miniature training evaluation	.48
job tryout	.44
biographical questionnaire	.40
structured interview	.25
academic performance	.20
unstructured interview	.20
reference check	.14
training and experience rating	.13
personality/ interest test	.10

(Pearlman, 1985).

1.02 PSYCHOLOGICAL TESTS

Psychological tests have been defined by Hull (1928) as being the measurement of some phase of a carefully chosen sample of an individual behavior.

"Traditionally the function of psychological tests has been to measure differences between individuals or between the reactions of the same individuals on different occasions" (Anastasi, 1982, p. 3).

Development of psychological tests began in the nineteenth century with the need to identify the mentally retarded. Now psychological tests can be found in a number of settings including clinical, educational, research, vocational, and industrial and are used in the solution of a wide range of problems.

There are three main categories of psychological tests used for personnel selection;

1. Tests of human abilities. eg cognitive abilities, psychomotor skills.

2. Tests of achievement of job specific abilities. eg operating machinery.

3. Personality and interest tests.

(McCormick and Tiffin, 1974).

Psychological tests are used in personnel selection because there is a commonly held belief that different characteristics found in different people will influence their ability to perform certain jobs. Hence psychological tests are used in attempts to identify individual characteristics that may increase the probability of success in that particular job. However the use of psychological tests for personnel selection is one of the most controversial subjects in the field of Industrial and Organisational psychology.

1.03 PERSONALITY TESTS.

Personality tests contrast most other tests used for personnel selection. Personality tests generally utilise a multivariate procedure where people are compared on a number of personality or interest dimensions to a norm. Such norms are part of the personality test package and some tests have norm data available which have been calculated for a number of different countries.

There are hundreds of personality tests available which fall into one of three main categories;

A: PROJECTIVE TESTS.

These tests involve the subject being presented with ambiguous stimulus such as inkblot pictures and being asked to interpret what they think they see. The respondent is assumed to project their personality onto the ambiguous stimulus

and by reporting what they experience reveal the kind of person they are. Conclusions are made by an expert who subjectively interprets the subject's responses. Two of the more common tests used are the Rorschach Inkblot Technique and the Holtzman Inkblot Technique (Anastasi, 1982). Reilly and Chao (1982) reviewed six studies on projective tests and found a mean validity of 0.28 and concluded that projective tests are not a viable alternative to other selection methods. Such tests are used generally in a clinical rather than occupational setting.

B: INTEREST TESTS.

Interest tests were designed to assess people's interests in different fields of work, and hence are commonly used for vocational guidance. They tap into the strength of the individuals interest in such things as hobbies, recreation, sport and jobs. The theory behind them is that if people are interested in what they are doing they will be better at it. Three of the more common interest inventories are the Strong Vocational Interest Blank (SVIB), the Kuder Preference Record, and the Vocational Preference Inventory (Anastasi, 1982).

C: PENCIL AND PAPER PERSONALITY QUESTIONNAIRES.

These questionnaires contain items relating to behaviours, attitudes and feelings, to which the subject responds. These tests generally are forced choice format and scoring is done by tallying up responses and comparing them to existing data and

norms.

Such tests include surveys, scales, profiles and schedules and are developed through internal consistency analysis. Some of the more common pencil and paper personality questionnaires used for personnel selection are the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway and McKinley 1967), the California Psychological Inventory (CPI; Mergargee 1972), the Sixteen Personality Factor Questionnaire (16PF; Cattell, Eber and Tatsuoka 1980), the Edwards Personal Preference Schedule (EPPS; Edwards 1957) and the Guildford-Zimmerman Temperament Survey (Guildford and Zimmerman 1956)(George 1986). The MMPI and the CPI are similar because the CPI was essentially derived from the MMPI. The MMPI however was intended for an 'abnormal' population whereas the CPI was not.

Blinkhorn and Johnson (1990) state that in recent years there has been a dramatic growth in the use of such tests for recruitment and selection and claim that about 50 percent of companies in the United Kingdom use them at some point in their selection assessment processes. The most common personality tests among British occupational psychologists is the 16PF (Tyler and Miller, 1986). This test is commonly used in New Zealand (Hesketh, 1974) and Australia (Smith, Dowling and Barry, 1985). Ryan and Sackett (1987) found that in the United States of America the most commonly used tests for individual assessment are the 16PF, the Guildford-Zimmerman Temperament Survey, the CPI, the MMPI, the Myers Briggs Type Indicator and the Edwards Personal Preference Test.

1.04 RELIABILITIES AND VALIDITIES.

Ghiselli and Barthol (1953) carried out the first major review of personality tests used for personnel selection and found a wide range of validities; negative, positive, large and small. Guion and Gottier (1965) carried out a more comprehensive review some years later after a large number of new personality measures were developed such as factorial inventories, new forms of projective measurements and forced choice inventories (Guion and Gottier, 1965). In their review Guion and Gottier included interest inventories because both interest inventories and pencil and paper personality tests have a common role of measuring motivation. Guion and Gottier's main conclusion was

"that taken as a whole, there is no generalizable evidence that personality measures can be recommended as good or practical tools for employee selection...[and that]...the best that can be said is that in some situations, for some purposes, some personality measures can offer helpful predictions" (Guion and Gottier, 1965, p. 159).

Since these two reviews, quantitative review techniques have emerged, and recent literature has been dominated by meta-analysis. Meta-analysis is a method of validity generalisation and

" although controversial, is a very powerful tool in modern research " (Ng, 1990, p. 12).

Early research on meta-analysis began before the 1960s, with much development occurring in the 1970s. Hunter, Schmidt and Jackson (1982) did much to improve meta-analytical techniques by emphasising the importance of correcting results for artifacts and methodological problems.

Schmitt, Gooding, Noe and Kirsh (1984) conducted a meta-analysis on predictors of job performance and reported an overall validity coefficient of 0.15 for personality tests, including measures of interest. Schmitt et al (1984) however did not account for major artifacts such as error of measurement on the dependent and independent variables.

Hunter and Hunter (1984) carried out a meta-analysis which did not include Schmitt et als (1984) findings due to the fact they were being performed concurrently. Inconsistencies were found between Hunter and Hunter's (1984) and Schmitt et als (1984) results. Hunter and Hirsh (1987) later concluded that these inconsistencies were due to an absence of corrections for error of measurement and attenuation. Hunter and Hirsh (1987) recalculated Schmitt et als. (1984) data using Hunter and Hunter's (1984) meta-analytic formulas which correct these errors and found the overall validity coefficients of personality tests used for predicting performance increased by 31 %, used for training criteria increased by 13 %, and used for status change and promotion increased by 3 %. The average overall validity of personality tests as calculated by Hunter and Hirsh (1987) was 0.19.

Ng and Smith (1991) updated a meta-analysis of personality tests and included in their study studies from seven journals dating from 1953 to 1990. Unpublished studies were also sought. Their results showed that personality tests had an overall validity of 0.25 and the variance due to sampling error was 10 %.

However much controversy still arises on the issue of using personality tests for personnel selection. Ng and Smith's (1990) conclusion that

" from the Guion and Gottier (1985) review to the present study 25 years later, non-work related personality tests used as selection tools should be treated with caution" (p.181)

appears to be a fair statement.

1.05 WORK RELATED PERSONALITY TESTS.

Blinkhorn and Johnson (1990) claim that such results as those mentioned above and results from many other studies may be due to

" details of construction leading to poorer psychometric properties; the clinical orientation of many of the test items; the scarcity of well conducted validity studies; and the use of discredited theoretical approaches to the structure of personality " (p.671).

Other explanations have been that people use deception and/or answer in a more socially desirable way. This issue is discussed later in 1.07.

Research is now focusing on the relationship between the content of the personality test and the content of the job. It is believed that if the questions in the personality test were related to a particular job, the validity of that test (when used in relation to the job it was devised for) would be better than those of other personality tests under similar circumstances.

It is likely that different personality variables are more relevant for different occupations. Day and Silverman (1989) give the example of accountants compared to firefighters, and in their experiment found three personality scales (work orientation, ascendancy and interpersonal orientation) showed promise for predicting job performance of accountants. From this they concluded

" that there is a place for personality measures when they are properly matched to a particular occupation and organisation " (Day and Silverman, 1989, p. 35).

In New Zealand a test called the Sales Performance Indicator (SPI (Anon, 1987)) has been developed and is used to predict the suitability of people for sales work. Three traits are measured; the persons drive to achieve results, their ability to relate to others, and their preferences in terms of independence and self determination in their work. Validity studies have been performed on the SPI which have found validities ranging between 0.5 and 0.6. However not enough

information is given on how such figures were reached, and therefore they should be treated with caution.

The SPI claims to be useful for selecting sales people regardless of the product being sold or the organisation selling the product.

Tests such as these are encouraging but more research is needed in this area.

1.06 SELF ASSESSMENT IN PERSONNEL SELECTION.

The psychological research of self is extensive, providing a wide diversity of perspectives and volumes of literature. Early researchers considered the assessment of oneself was influenced by environmental influences, unidirectional effects and internal states. Present research is focusing more on self assessment as a part of an ongoing process and as an agent reflecting social context. Self assessment is seen as a active agent in that it can regulate behaviour through observation, judgement and reaction (Bandura, 1982b). These will be influenced by a large number of factors such as what is focused on, its functional relevance, value orientations, referential information, importance and difficulty of activity, the degree to which results can be attributed to oneself, incentitives, self consequences and motivational properties among other things (Bandura 1977b, 1982b, 1978).

Self assessments have been used in a number of practical settings;

a. evaluation in current performance for administration purposes.

b. identification of training needs eg career planning.

c. criterion measurement in applied research.

d. personnel selection.

e. construct measurement in basic research.

(Thornton, 1980, p. 264).

It is accepted within the literature on self that people are able to know and assess themselves (Bandura, 1978; Bem, 1972; Mischel, 1972; Wicklund 1975). Individuals have more data about themselves than does an evaluator and they may process this data in ways that may lead to greater accuracy for prediction (Schrader and Osburn, 1977). Empirical research too has found that individuals are more sensitive than external observers to situational determinants of their behaviour and are less likely to over or under attribute the outcome of their actions to dispositional factors (Fox and Dinur, 1988).

In most situations where assessments are being performed the individual being assessed has as much to gain or lose from decisions made due to this assessment. It therefore seems fair that if the individual is capable of accurate self assessment then these should be encouraged and incorporated into any decision making. Furthermore self assessments are more likely to be economical in terms of time and money than other assessment methods.

There is increasing research on self assessment used in a number of situations with some studies in favour of self assessment and others not. In a review of self appraisals of work performance Thornton (1980) found that self appraisals tend to show more leniency, less variability and less discriminant validity. However the

"inescapable conclusion from this review is that individuals have a significantly different view of their own job performance than that held by other people " (p. 268).

Schrader and Osburn (1977) reviewed the accuracy of self-prediction and judgements by others in psychological assessment and concluded that self assessments are at least as predictive as other assessment methods for intellectual achievement, vocational choice, job performance, therapy outcome, adjustment following hospitalisation and peer ratings.

In their review on 55 studies, in which self evaluations of ability were compared with measures of performance, Mabe and West (1982) found that under certain measurement conditions self assessment of ability may closely correspond to performance on criterion measures. A number of factors were found to influence the accuracy of these self assessments. High intelligence, high achievement status and internal locus of control were associated with more accurate evaluations as were previous experience at self assessment, the rater's expectation that their self assessment would be validated against other measures, a guarantee of anonymity of the self evaluation and self evaluation instructions emphasising comparison

with others.

Primoff (1980) states that

" self assessments are a useful tool for selection when people are asked to rate themselves on the specific aspects of work behaviour which constitute successful job performance " (p. 283).

Factors that affect the relationship between self assessment and other variables were found to be the extent of understanding between the self raters and those who prepare the test of the behaviour statements and the extent to which self raters have a common base for their ratings. Primoff (1980) claims that self assessments increase accuracy and fairness in examining when combined with other information such as achievements and test scores.

Most studies of self assessment have been contrived and have not involved actual situations of selection or decision making (Schrader and Osburn, 1977). Fox and Dinur (1988) evaluated the validity of self assessment in a natural setting. The subjects were being screened for a prestigious Military Course over nine days and self assessment measurements were taken mainly on personality attributes rather than abilities. Results indicated that self assessments were low but significant for predicting success over a two year training period and partially support the notion of self assessment being a valuable tool for personnel selection.

George and Smith (1990) looked at self assessment as a predictive tool for personnel selection by requesting applicants to self assess their abilities on seven work sample exercises before and after an assessment centre. Applicants who became employed also provided self assessments of job performance after one year on the job. The results found that self assessment showed no predictive utility as a predictor of job performance and that the assessment centre reduced the self assessment ratings on performance much like realistic job previews do. Furthermore no moderator effects for age, gender, intelligence, social desirability, full time work experience and education were found to effect the relationship between self assessment and organisational estimates of work performance.

1.07 SELF ASSESSMENT AS A PREDICTIVE TOOL.

There is still limited research on self assessment as a predictive tool for personnel selection. De Nisi and Shaw (1977) highlight the problem of self assessment in a real life situation by asking

" if people can do it, will they do it accurately in real life ? " (p. 641).

Wicklund and Gollwitzer (1983) claim the relationship between what people say they will do and what they actually do is often weak. Studies have found a lack of consistency between expressed attitudes and actual behaviour (Fazio and Zanna, 1978; Regan and Fazio, 1977; Wicker, 1979; Zanna, Olson and Fazio, 1980). However research on trait approaches to personality and attitude-behaviour

consistency generally has not examined the predictive value of self assessment relative to other less direct means of predicting behaviour.

1.08 PROBLEMS WITH SELF ASSESSMENT.

Skeptism about self assessment used in an applied setting has been caused mostly by the following factors;

a) cognitive and motivational processes.

b) unconscious feelings.

c) response bias.

Cognitive and Motivational Bias.

Situations that involve the allocation of scarce resources such as jobs, promotions or entry to professional training particularly elicit skeptism when self assessment is used. It is assumed that individuals' motivation will result in distortion of their abilities and therefore predictive validity of their self assessment relative to other criteria will be low. Scott and Johnson (1972) however found that self assessment obtained in situations designed to make candid responses potentially costly have still been more accurate than other predictors.

Schrader and Osburn (1977) claim

" the motivation to present oneself in a favourable light should invalidate self-assessments only when it is stronger than the desire to be accurate" (p. 347).

Efforts can be made to develop collegial and collaborative relationships between the assessor and assessed so that it becomes easier to communicate that accurate and valid self reporting is the socially desirable response (Schrader and Osburn, 1977).

Unconscious Feelings.

The argument that people are inaccurate at appraising themselves because of underlying unconscious feelings comes from a psychoanalytic perspective. This theory assumes that people are usually unaware of some of their most important feelings and that these feelings strongly influence people's responses. This view has led professionals to question the validity of self assessment and instead of focusing on the obvious content of the responses focus instead on what underlying unconscious feelings may be apparent. (Schrader and Osburn, 1977).

Response Bias.

The knowledge of the self is a product of social interaction. This is illustrated by there being much research on social comparison. Earlier research focused on

individual knowledge derived from internal states, but more recently the importance of social context has been emphasised (Markus and Wurf, 1987).

George (1988) found few New Zealand practioners accepted self assessment as a useful selection method. The reluctance of using self assessment in an applied setting stems from the belief that responses are biased due to self enhancement desires. Accuracy of ones self assessment may be influenced by self motives being served and on the immediate social situation. For example Wicklund and Gollwitzer (1983) claim that to use self assessment as a selection tool problems arise in that there is a tendency to use self assessment to compensate for any weaknesses by the applicant. Wicklund and Gollwitzer have termed this behaviour 'anti-validity'.

Unfortunately there is little research on the validity of self assessment in work related situations such as personnel selection. Some research findings have indicated response bias in the direction of social desirability for personality inventories (Elliot, 1976; Elliot, 1981; Michaelis and Eysenck, 1971; Thornton and Gierasch, 1980), interest measures (Zalinski and Abrahams, 1979) and skills and abilities (Ash, 1980).

However, recently Hough, Dunnette, Eaton, Kamp and McCloy (1990) used personality scales that had obvious items that could be easily distorted and found that

"intentional distortion of self descriptions in an overly desirable way does not

appear to be a serious problem. Applicants did not appear to distort their self descriptions, and correlations with job performance were not attenuated by such distortion" (p. 593).

Studies that have found no inflation bias of self assessment of abilities and skills were performed when applicants were aware that their self assessments would not result in decisions affecting their potential employment (De Nisi and Shaw, 1977; Levine, Flory and Ash, 1977).

Although there is debate over inflation bias in job applications, opinions are mixed on the effect of such bias on selection validity. Elliot (1981) considers it healthy to respond in a manner compatible with what is expected and it has been noted that inflation bias may not reduce examination validity when personality measures are used. Kreidt and Dawson (1961) found a positive relationship between socially desirable responses and clerical job performance and Ruch and Ruch (1967) found that response set was the main predictor of job performance for sales people.

1.09 FACTORS AFFECTING SELF ASSESSMENTS.

Few studies have found any moderating effects on the relationship between self reported and tested abilities for personality traits. In their review Mabe and West (1982) found high intelligence, high achievement status and internal locus of control were associated with more accurate self evaluations of ability. These

person variables and measurement conditions examined in Mabe and West's (1982) review do not comprise a comprehensive list of factors that may potentially influence the validity of self evaluation of ability.

De Nisi and Shaw (1977) found gender, intelligence, social desirability and self esteem to have no relationship between self reported and tested abilities. George and Smith (1990) looked at possible moderators such as age, gender, intelligence, social desirability, full-time work experience and education and found no consistent moderator effects.

However a number of potentially relevant factors have not been examined. Wicklund (1979) discovered that as people become more self aware they reduce discrepancies between self descriptions and actual behavior. It is possible that conditions increasing objective self-awareness could operate to increase the validity of self assessments (Mabe and West, 1982).

Primoff (1980) found correlations between self assessments and other variables are affected by the extent to which the self assessors have the same understanding of the scaled behaviour statements as those who prepare the test or examination and the extent to which self assessors have a common base for their ratings. Mabe and West (1982) claim that improving the validity of self assessment can be achieved by giving instructions to subjects that their self assessments would be validated against other information. Furthermore the subjects' previous experience at self assessment can improve validity and therefore more careful

structuring of these self assessment experiences may help elicit valid self assessments.

Schrader and Osburn (1977) have suggested that external evaluation might facilitate optimal judgements by;

a) being explicit to the self assessors what aspects of their behaviour are to be predicted and in what contexts. Questions asked should be specific rather than general and for personality inventories items should have face validity rather than be subtle (Holden and Jackson, 1979).

b) phrasing questions to maximise accuracy.

c) facilitating recall of relevant previous experience. Empirical research has found the validity of self predictions is increased if people have had experience in the situations they are being asked about. It is suggested that the evaluator assess the amount of experience the self assessor has had in these situations and also maximise their recall of those relevant previous experiences.

d) evaluating and encouraging the self assessors motivation for accuracy.

Hough, Dunnette, Eaton, Kamp and McCloy (1990) outline a number of strategies used to overcome intentional distortion.

a) The use of the forced choice format where the choices of responses are of equal social desirability.

b) The use of subtle items - items for which the underlying construct is not apparent. Research in this area however suggests that subtle items are often less valid than obvious items (Duff, 1965; McCall, 1958; Wiener, 1948,) and this may reduce scale validity. Furthermore, Hough et al (1990) claim that

" the assumption that subtle items are less susceptible to distortion has never been satisfactorily demonstrated" (p. 582).

c) Warning subjects about the consequences of distorting their responses and that detection methods and verification procedures exist. This approach has been found to be effective in reducing intentional distortion (Haymaker and Erwin, 1980 cited by Hough, Eaton, Dunnette, Kamp and McCloy 1990; Schrader and Osburn, 1977; Trent, Atwater and Abrahams, 1986 cited by Hough, Eaton, Dunnette, Kamp and McCloy 1990).

d) Develop scales or methods to identify suspicious self assessments. A number of personality inventories include response validity scales such as the California Psychological Inventory (CPI), the Minnesota Multiphasic Personality Inventory (MMPI) and the Multidimensional Personality Questionnaire (MPQ).

In their study Hough et al (1990) found that intentional distortion of self

assessment on personality measures was not a serious problem. However they offered the following approaches when using temperament measures in a selection setting;

" (a) use response validity scales to detect potentially inaccurate self-descriptions; (b) warn applicants that inaccurate descriptions will be detected; and (c) use additional or other information to make employment decisions about those people who are identified as providing inaccurate self-descriptions" (p. 594).

1.10 WHAT IS A PERSONALITY?

A number of definitions have been offered from different fields of psychology to describe what a personality is. Such definitions include;

"...the dynamic organisation within individuals of those psychophysical systems that determine his [or her] characteristic behaviour and thought" (Allport, 1961, p.28).

"...a person's unique pattern of traits" (Guilford, 1959,p.5).

"...the distinctive patterns of behaviour (including thoughts and emotions) that characterise each individual's adaptation to the situations of his or her life" (Mischel, 1976, p.2).

"Personality is a stable set of characteristics and tendencies that determine those commonalities and differences in the psychological behaviour (thoughts, feelings, and actions) of people that have continuity in time and that may or may not be easily understood in terms of the social and biological pressures of the immediate situation alone" (Maddi, 1972,p.9).

It is difficult to describe the ultimate measure of personality, particularly when an agreement on the definition of personality is not universally accepted. The goal in measuring personality is to obtain an approximate estimate of this ultimate criterion (ie an individual's personality) by selecting one or more actual criteria (such as a personality test), which are considered appropriate (Blum and Naylor, 1968). However, we generally can not measure the ultimate criterion, and therefore are left not knowing the degree of criterion relevance (ie the degree to which the actual criterion overlaps to the ultimate criterion).

The present study compares two actual criteria; individuals' personality test scores and their corresponding self assessment scores on the factors and scales of the personality tests, and attempts to evaluate the results in relation to personnel selection. There are three experiments which look at whether the subjects are accurately able to assess their personalities as measured by personality tests, and whether, when doing their self assessments, the subjects respond in a socially desirable way.

1.11 STATEMENT OF THE HYPOTHESES.

The present study attempts to evaluate self assessment on personality tests for personnel selection.

The hypothesis for the first two experiments is that people will be able to accurately assess their personalities as measured by personality tests. Whether subjects' previous experience at self assessments, confidence at the task, subject's gender and the order of which the subjects do the two tasks (ie the self assessment and the personality test) has any effect on the relationship between the subject's self assessment and personality test scores will be tested. It is hypothesised that of these possible moderators only previous experience at doing self assessments may increase the subjects' accuracy at assessing their personalities as measured by personality tests (ie higher correlations between the subjects' personality test scores and self assessment scores will be found for subjects who have had previous experience at self assessment). These hypotheses will be tested using two different personality tests; the California Psychological Inventory (CPI) and the Sixteen Personality Factor Questionnaire (16PF). These tests were chosen because they are two of the most commonly used personality tests used in New Zealand for personnel selection.

A third experiment will be performed to find if certain responses are considered to be more socially desirable. It is hypothesised that factors on the 16PF and scales on the CPI are not socially neutral. By this it is meant that one pole of each

factor or scale is more socially desirable than the opposite pole. The developers of the 16PF claim that each factor of the 16PF is socially neutral. However, Merrideth (1968) has shown that the poles of the 16PF are not equally attractive. Due to different techniques of test development the producers of the CPI do not make this same claim.

It is also hypothesised that the subjects when doing their self assessments will not respond in a socially desirable way. By this it is meant that the subjects' personality test scores and self assessment scores will not be significantly different and that subjects will not rate themselves, when doing their self assessment, towards the more attractive pole.

If these hypotheses prove to be valid these findings will cast further doubt on the use of personality tests for personnel selection. Instead if companies insist on getting personality data on applicants it would be cheaper and quicker for applicants to rate their own personalities. If these hypotheses are not supported by the results the question of which is a more accurate measure of a person's personality, the personality test scores or their self assessment scores ? will need to be confronted.

CHAPTER TWO

METHOD

Experiment One

2.01 SUBJECTS

The subjects were 57 volunteer students from Massey University, New Zealand. There were 30 (53%) males and 27 (47%) females. The age range was 18 to 39 years, with the average and modal ages being 21 years and 18 years respectively.

95 % classed themselves as New Zealand European while 5 % classed themselves as either Maori, Polynesian, or Asian.

The subjects were recruited by advertisements in the student newspaper, the University newsletter and by being approached in first year psychology classes and asked to volunteer as participants in the study. The subjects were told that they would get the opportunity to do a personality test, one that is commonly used in New Zealand for personnel selection.

2.02 MATERIALS

1. The Sixteen Personality Factor Questionnaire (16PF).

-16PF HISTORY-

Allport and Odbert (1936) identified 17,954 trait names relative to human behaviour. They reduced them to 4,504 traits which they labelled 'real traits'. Cattell used this as a starting point and reduced this list further to 171 traits by eliminating synonyms. Then through a cluster analysis of peer ratings he classed these traits into clusters called 'surface traits'. Using factor analysis Cattell eventually identified 20 distinct factors he called 'source traits'. Four of these source traits were dropped after further calculations. These sixteen remaining source traits are believed to be inherent ones underlying more manifest behaviour traits.

-16PF DESIGN-

There has been some criticism on the use of factor analysis for developing the 16PF. Francis (1972) clearly highlighted that factor analysis should only be used in psychological research when the number of factors is known. Francis (1972) also showed that factors could be found from intercorrelated random error. In an example he found loadings for three non-existent factors which were reproduced in three independent samples. Cattell however used factor analysis to uncover

relationships, and did so using repeated application of factor analysis. Bull (1974) concluded from his work with the 16PF that

" a factor analysis may indicate structure but in most designs it has nothing to do with the validity of a test with respect to applied psychology. The measure of science is its ability to predict, and we must demonstrate a relationship between predictor and criterion variables in order to justify the continued use of a test, in anything other than a research situation " (p.14).

-RELIABILITY AND VALIDITY-

Research has shown the scales as not being very reliable across the different forms of the test, with the median coefficient for the 16 scales being 0.44 (Cattell, 1972). This raises the issue of stability of the factor structure, especially when used among different populations and for identification of factors across different methods. Furthermore Zuckerman (1985) has claimed that

"specific factor weightings are meaningless without replication...[and that]...very little replication is reported in the handbook" (p. 1393).

Peterson (1965) found no acceptable replication across age and Eysenck, White and Souief (1969) factor analysed Cattell's 16PF items and found that the resultant primary factors were not readily identifiable with those of Cattell and were not replaceable for males to females. Studies carried out in England (Eysenck, 1969),

Canada (Howarth and Brown, 1971) and New Zealand (Adcock, Adcock and Walkey, 1971) have found the factors lack consistency in different countries. Cattell however claims that there is a wide applicability across diverse areas of practice and that there is cross-cultural trait stability. In fact there are now standardisations of the 16PF in effective translations in over 15 different countries.

The number of items per factor is either 10 or 13. In a study using New Zealand University students Adcock (1974) found many items were measuring related factors and not those for which they were designed. She also found five factors (N, Q2, Q1, L and M) failed to achieve better than chance discrimination. The most promising factors were I, H, F and G but these still had faults in areas such as more significant correlations outside the factor cluster.

-USERS-

A problem with the 16PF is that its limits of applicability are not clearly defined. It was developed as a research instrument for assessing source traits. Cattell claims

"...the traits covered by the 16PF are those centrally concerned in all these fields [clinical, industrial and educational] of application. Thus the 16PF is designed as an all purpose instrument bringing to applied psychology the concepts central to general personality theory " (Cattell, 1972, p. 10).

The 16PF has been selected as an instrument for the study of personality in a multiple of settings. Such areas include

" personality and motivational factors related to potential aviation accidents among U.S Naval Academy graduates to the study of personality differences between U.S. Olympic national and nonnational swimmers.....[and] personality relationships to such areas as cancer treatment, cross-cultural acclimation, teachers of the emotionally disturbed, juvenile delinquency, religion, child abuse, imminent death anxiety, air traffic control, nutrition behaviour, controlled substance abuse, long-term effects of concentration camp internment, attitudes towards variable physiological functioning, the screening of seminarians for the priesthood, and child adoption" (Wholeben, 1986 p. 325).

It has also been used for personality research, pre-marital counselling, law enforcement officer selection and clinical problem cases. More recently the 16PF has been used increasingly for career guidance, vocational exploration and occupational testing.

Butcher (1985) suggests however that it is too much to expect one test to provide substantive personality descriptions that span such a broad field. Butcher (1985) claims that the 16PF is valuable for personnel selection, guidance counselling and personality research but because it is designed for a normative population it should not be used as a clinical instrument. Zuckerman (1985) however points out that for tests used as instruments of research construct validity is crucial, but when

used in an applied setting, criterion related validity is more important. The 16PF was designed as a research instrument yet it makes claims as to applicability in problems of selection and evaluation.

-DESCRIPTION OF THE 16PF-

The 16PF is constructed in five forms; A,B,C,D, and E with each measuring the same 16 personality dimensions. These are;

Warmth (factor A): reserved, critical, detached, aloof versus outgoing, warmhearted, easy-going, participating.

Intelligence (factor B): less intelligent, concrete thinking versus more intelligent, abstract-thinking, bright.

Emotional stability (factor C): affected by feelings, emotionally less stable, easily upset versus emotionally stable, faces reality, calm, mature.

Dominance (factor E): humble, mild, accommodating, conforming versus assertive, aggressive, stubborn, competitive.

Impulsivity (factor f): sober, prudent, serious, taciturn versus happy-go-lucky, impulsively lively, gay, enthusiastic.

Conformity (factor G): expedient, disregards rules, feels few obligations versus conscientious, persevering, staid, moralistic.

Boldness (factor H): shy, restrained, timid, threat-sensitive versus venturesome, socially bold, uninhibited, spontaneous.

Sensitivity (factor I): tough-minded, self-reliant, realistic, no-nonsense versus tender-minded, clinging, over-protected, sensitive.

Suspiciousness (factor L): trusting, adaptable, free of jealousy, easy to get along with versus suspicious, self-opinionated, hard to fool.

Imagination (factor M): practical, careful, conventional, regulated by external realities, proper versus imaginative, wrapped up in inner urgencies, careless of practical matters, bohemian.

Shrewdness (factor N): forthright, natural, artless, unpretentious versus shrewd, calculating, worldly, penetrating.

Insecurity (factor O): self-assured, confident, serene versus apprehensive, self-reproaching, worrying, troubled.

Radicalism (factor Q1): conservative, respecting established ideas, tolerant of traditional difficulties versus experimenting, liberal, analytical, free-thinking.

Self-sufficiency (factor Q2): group-dependent, a 'joiner' and sound follower versus self-sufficient, prefers own decisions, resourceful.

Self-discipline (factor Q3): undisciplined self-conflict, follows own urges, careless of protocol versus controlled, socially precise, following self-image.

Tension (factor Q4): relaxed, tranquil, unfrustrated versus tense, frustrated, driven, overwrought.

There are also four second order factors based on factor loadings from each of the 16 profile aspects of human behaviour. These are extraversion, anxiety, tough poise and independence.

-ADMINISTRATION OF THE 16PF-

Administration of each form is 30 - 60 minutes depending on the form. The separate forms vary in readability requirements and the number of items vary for each test with A and B having 187, C and D having 105, and E having 108 (see table two).

All forms are intended for use on people 16 years and older and are in the form of forced-option format where there are three options to chose from. The subjects respond to the test items by putting their answer on a prepared answer sheet. Test administration is usually untimed. The administer then calculates the raw scores

Table Two

The Five Forms of The 16PF

Form	No. of items	Use	Approx. time (minutes)
A	187	Ordinary newspaper-literate adults	50
B	187	Ordinary newspaper-literate adults	50
C	105	Vocabulary demand somewhat less than for A and B -average adults	30-40
D	105	Vocabulary demand somewhat less than for A and B -average adults	30-40
E	128	Very simple vocabulary -low-literate groups	30-40

(Cattell, Eber and Tatsuoka, 1970, p.3)

for each of the 16 factors by the use of templates which are then translated to sten scores. These are based on normative tables supplied with the instrument. The sten scores are then plotted on to a profile form which shows the subject's scores on a scale of one to ten on each of the 16 factors. To calculate the sten scores for the four second order factors regression equations are used. This involves using the 16 sten scores as weighted factor loadings.

It is recommended that tests be used in pairs to increase their precision. Realistically however users do not have the time to do this.

Interpretation of the results is established by the studying the profile form or through computer generated interpretive reports available from the publisher.

-JUSTIFICATION-

For the present experiment form A was used because the subject population were University students. Form B was not used in conjunction because users rarely administer both due to the time involved. Factor B (intelligence) was included in the present experiment out of interest, even though other researchers have excluded it from their calculations. The Handbook suggests this factor not be interpreted or used in isolation and that a supplementary intelligence test be used in conjunction. The second order factors were not included because they are not present on the profile form

2. The Profile Form for the 16PF (see appendix A).

The profile form is generally used as part of the administrative side of the 16PF. It is used by the test marker to plot the subject's personality scores on the sixteen factors. However, in the present experiment it was also used by the subjects to do a self assessment and plot themselves as they see their personalities.

3. The General Questions Form (GQF) (see appendix B).

The GQF was designed by the experimenter to measure subjects previous experience at self assessment and their confidence at self assessing their personalities as well as finding out general information about the subjects such as age, gender and ethnic group.

2.03 ETHICAL CONSIDERATIONS

1. Informed consent.

Informed consent was obtained in two ways. Firstly subjects agreed to participate and secondly they signed a consent form (see appendix C).

2. Debriefing.

Subjects were debriefed once the experiment was completed. A written and

verbal explanation was given and discussion was encouraged.

The procedure and ethical considerations were discussed and approved by the Massey University Ethics Committee.

2.04 PROCEDURE

Subjects placed themselves into one of two experimental conditions, depending on which time slots suited their timetables. The experiment had two parts to it spaced one week apart. There were also two conditions, condition A and condition B. In condition A subjects did the personality test first and the self assessment one week later. In condition B subjects did the self assessment first and the personality test one week later. This was done to test for any ordering effects.

CONDITION A.

Part One.

Subjects in condition A (19 males and 22 females) were administered the 16PF in a classroom. They were given code numbers to use for identification instead of using their names. This was done to preserve anonymity. They were also asked to place their date of birth on their answer sheets as a back up for identification in case subjects forgot their code number the following week when doing the second

part of the experiment.

Before doing the personality test subjects filled in a consent form and were then told to read the instructions on the front page of the 16PF carefully. They were reminded that they were free to drop out of the experiment at any time without giving any explanation.

The subjects could start when they were ready. The experimenter remained seated in the room with the subjects throughout the duration of the test, which took on average about 45 to 55 minutes to complete. On completion the subjects were reminded that the second part to the experiment was a week away and that the experimenter would call to remind them the night before that date. They were also told to leave their names and experimental code numbers on a drawn up piece of paper on the experimenter's desk if they were interested in getting their individual results back.

Part two.

Part two was completed in the same classroom at the same time of day a week later. The subjects were given the GQF to complete. The subjects were then asked to rate themselves as they believe themselves to be on the 16PF profile form.

CONDITION B.

The subjects in condition B were 11 males and 5 females. Conditions were identical to those in condition A except the subjects in condition B did part two, the GQF and self assessment first, and part one, the personality test a week later. For both condition A and B the consent form was completed before the experiments commenced. This meant that the subjects in condition A did this before the personality test and subjects in condition B did this before the GQF and their self assessment on the 16PF profile form.

2.05 DEBRIEFING

Debriefing occurred immediately after the subjects had completed the second part of the experiment. The debriefing consisted of a verbal discussion and outline of the experimental aim and procedure and a written verification of these. The subjects were encouraged to discuss any points or ask any questions.

2.06 FEEDBACK

Feedback was given approximately three weeks later when subjects collected their individual results. The subjects were also given the overall results for the experiment that had been calculated that far.

CHAPTER THREE

RESULTS

Experiment One

Subjects were generally confident at rating their personalities on the profile form. On a scale of one to ten (with ten being extremely confident) the mean and modal responses were 8.3 and eight respectively, with no subjects rating their confidence at the task below five.

90 percent of subjects had no previous experience at self assessment of this nature and therefore this variable was excluded from further analysis.

Sixteen factorial analysis of variance (ANOVAs) were calculated to test whether condition and gender had any influence on subjects' accuracy at rating their personalities as measured by the 16PF. The dependent variables were the total of the differences between each subject's test score and self assessment score for each factor. Hence DIFFA equalsΣ(TESTA - SELFA). Two tailed correlations were performed across these 'DIFF' values and significant correlations were found to be few. Therefore factorial ANOVAs were performed rather than MANOVAs. Results from the ANOVAs were found to be non-significant ($p < 0.003$ ie .05/16 where 16 equals the number of ANOVAs calculated) indicating that gender and

condition had no influence on the subjects' accuracy at rating their personalities as measured by the 16PF.

The hypothesis that subjects are able to assess their personalities as measured by the 16PF was tested by calculating two tailed correlations ($p < 0.01$) between the test scores and the self assessment scores for each of the 16 factors. Of the sixteen correlations five were found to be significant; H, O, G, F, and E. Scatterplots were run for the 16PF test scores and the self assessment scores for the sixteen factors (see appendix D1 - D16) and means and standard deviations were calculated (see table three) to aid in interpretation of the correlations.

Factor H : Boldness; shy, restrained, timid, threat-sensitive versus venturesome, socially bold, uninhibited, spontaneous.

Factor H found the strongest relationship between the 16PF test scores and the subjects' self assessment scores with the correlation being 0.6524 ($p= 0.000$). The scatterplot shows a positive linear relationship with the data for both measures skewed in favour of higher scores. The mean 16PF test scores and self assessment scores were 6.77 and 6.74 respectively.

Factor O : Insecurity; self-assured, confident, serene versus apprehensive, self-reproaching, worrying, troubled.

Factor O yielded a high correlation between the 16PF test scores and the subjects' self assessment scores of 0.5463 ($p= 0.000$). The scatterplot indicates a positive linear relationship with the mean 16PF and self assessment scores being 5.84 and

4.53 respectively.

Factor G : Conformity; expedient, disregards rules, feels few obligations versus conscientious, persevering, staid, moralistic.

Factor G yielded the third strongest correlation between 16PF test scores and subjects' self assessment scores of 0.3735 (p= 0.004). The scatterplot and the group means (4.14 for the 16PF test scores and 6.97 for the self assessment scores) indicate that subjects' self assessments on this factor are significantly higher than the test calculated.

Factor F : Impulsivity; sober, prudent, serious, taciturn versus happy-go-lucky, impulsively lively, gay, enthusiastic.

The 16PF test scores and the self assessment scores correlated 0.3616 (p= 0.006) for this factor. Like factor H the scatterplot indicates a positive linear relationship with data from both measures skewed in favour of higher scores. The mean scores for the 16PF test scores and the self assessment scores on factor F are 7.12 and 7.02 respectively.

Factor E : Dominance; humble, mild, accommodating, conforming versus assertive, aggressive, stubborn, competitive.

The correlation between the 16PF test scores and the self assessment scores for factor E was 0.3549 (p= 0.007). Like factor H and factor F the scatterplot indicates a positive linear relationship with data from both measures skewed in

Table Three

Correlations between the 16PF test scores and corresponding self assessment scores and means and standard deviations for the 16PF test scores, self assessment scores and scores of the difference between the two.

16PF Scale	corr of test with self	p value	test score mean (n=57)	test score st.dev.	self assess mean (n=57)	self assess st.dev	diff scores	diff st.dev
H	0.6524	0.000	6.77	1.94	6.74	1.99	0.04	1.64
O	0.5463	0.000	5.84	2.06	4.53	1.90	1.32	1.89
G	0.3735	0.004	4.14	1.95	6.97	1.63	-2.83	2.02
F	0.3616	0.006	7.12	1.73	7.02	1.56	0.11	1.87
E	0.3549	0.007	6.88	1.87	6.81	1.47	0.07	1.93
A	0.3147	0.017	5.18	1.99	7.35	1.61	-2.18	2.13
Q4	0.2594	0.051	5.74	1.54	4.63	2.08	1.11	2.24
Q1	0.2572	0.053	6.54	1.99	6.79	1.97	-0.25	2.42
C	0.2102	0.117	5.42	1.89	6.81	1.83	-1.39	2.34
B	0.2042	0.128	6.07	2.27	7.58	1.07	-1.51	2.30
Q3	0.2013	0.133	5.72	1.84	6.40	1.71	-0.68	2.25
M	0.1580	0.240	5.72	1.79	5.21	1.98	0.51	2.45
I	0.1511	0.262	5.74	2.26	4.77	1.91	0.97	0.36
Q2	0.1076	0.425	6.09	2.02	6.84	1.79	-0.75	2.55
L	0.0638	0.638	6.39	1.95	4.32	1.87	2.07	2.62
N	-0.0252	0.852	3.60	1.95	4.98	1.96	-1.39	2.80

favour of higher scores. The mean scores for the 16PF test scores and the self assessment scores on factor F are 6.88 and 6.81 respectively.

Factor A : Warmth; reserved, critical, detached, aloof versus outgoing, warmhearted, easy-going, participating.

factor Q4 : Tension; relaxed, tranquil, unfrustrated versus tense, frustrated, driven, overwrought.

factor Q1 : Radicalism; conservative, respecting established ideas, tolerant of traditional difficulties versus experimenting, liberal, analytical, free-thinking.

Factor A, factor Q4 and factor Q1 show correlations of 0.3147 (p= 0.017), 0.2594 (p= 0.051) and 0.2572 (p= 0.053) respectively. For factor A the mean 16PF test score was 5.18 with the mean self assessment score being 7.35. The scatterplot indicates this as being a common finding for this factor with few subjects having scored or assessed themselves much differently. The scatterplot for factor Q4 indicates a positive linear relationship but with much of the data clustered centrally. The mean 16PF score is 5.74 and the mean self assessment score is 4.63. A linear relationship is not easily apparent for factor Q1. The majority of subjects scored above average and rated themselves above average on this factor with the mean 16PF scores and self assessment scores being 6.54 and 7.71 respectively.

Factor C : Emotional stability; affected by feelings, emotionally less stable, easily upset versus emotionally stable, faces reality, calm, mature.

Factor B : Intelligence; less intelligent, concrete thinking versus more intelligent,

abstract-thinking, bright.

Factor Q3 : Self-discipline; undisciplined self-conflict, follows own urges, careless of protocol versus controlled, socially precise, following self-image.

Factor M : Imagination; practical, careful, conventional, regulated by external realities, proper versus imaginative, wrapped up in inner urgencies, careless of practical matters, bohemian.

Factor I : Sensitivity; tough-minded, self-reliant, realistic, no-nonsense versus tender-minded, clinging, over-protected, sensitive.

Factor Q2 : Self-sufficiency; group-dependent, a 'joiner' and sound follower versus self-sufficient, prefers own decisions, resourceful.

Factors C, B, Q3, M, I and Q2 did not have significant correlations between subjects' 16PF test scores and subjects' self assessment scores. The correlations respectively are 0.2102 (p= 0.117), 0.2042 (p= 0.128), 0.2013 (p= 0.133), 0.1580 (p= 0.240), 0.1511 (p= 0.262) and 0.1076 (p= 0.425). The scatterplots generally indicate random data. No subjects rated themselves below average on factor B.

Factor L : Suspiciousness; trusting, adaptable, free of jealousy, easy to get along with versus suspicious, self-opinionated, hard to fool.

Factor N : Shrewdness; forthright, natural, artless, unpretentious versus shrewd, calculating, worldly, penetrating.

The lowest correlations between the 16PF test scores and the subjects' self assessment scores were found for these two factors. The correlation for factor L is 0.0638 (p= 0.638) and for factor N is -0.0252 (p= 0.852). These are below

chance probability. The scatterplots for these two factors indicate no linear relationships.

3.01 CONCLUSION

The hypothesis that gender and condition would not affect subjects' ability to assess their personalities as measured by the 16PF was supported. However the hypothesis that subjects would be able to assess their personalities as measured by the 16PF was only partially supported. Five of the sixteen factors (factors H, O, G, F and E) showed this to be true but the remaining factors did not. Factors L and N were found to have lower than chance probability of doing this accurately. The question of which is a better measure of one's personality will be discussed in chapter eight.

CHAPTER FOUR

METHOD

Experiment Two.

4.01 SUBJECTS

The subjects were 46 volunteer students from Massey University, New Zealand. There were 25 (54%) males and 21 (46%) females. The age range was 18 to 42 years, with the average and modal ages being 23 years and 19 years respectively.

87% classed themselves as New Zealand European while the other 13% classed themselves as either Maori, Polynesian, Asian or Lebanese.

The subjects were recruited by advertisements in the student newspaper, the University newsletter and by being approached in first year psychology classes and asked to volunteer as participants in the study. The subjects were told that they would get the opportunity to do a personality test, one that is commonly used in New Zealand for personnel selection.

4.02 MATERIALS

1. The California Psychological Inventory (CPI).

-DEVELOPMENT OF THE CPI-

The CPI was derived from the Minnesota Multiphasic Personality Inventory (MMPI) but was intended for use with a 'normal' population. The purpose of the CPI scales as stated by the author is

" to forecast what a person will say or do under defined conditions, and to identify individuals who will be described in characteristic ways by others who know them well or who observe their behaviour in particular contexts".

Furthermore the concepts chosen for assessment are

" folk concepts - aspects and attributes of interpersonal behaviour that are to be found in all cultures and societies, and that possess a direct and integral relationship to all forms of social interaction" (Gough, 1968).

The CPI scales do not purport to measure personality traits. Gough (1968) has grouped scales for

"convenience...bringing together those having related implications" (Gough, 1968)

The underlying logic is interpretational and not factorial, as is the 16PF. There is no major personality theory underlying the selection of the constructs. Baucom (1985) claims that the

"absence of an underlying personality theory is neither an inherent strength or weakness of the test", (p. 251)

and that Gough has chosen one of many methods to construct scales.

Eysenck (1985) however claims that Gough has tried to group scales that correlate highly without using correlational analysis.

Gough emphasises that the scales were not intended to define traits and therefore validation should be measured by the degree to which the test can forecast behaviour and identify individuals who are perceived in characteristic ways.

Table four shows the intercorrelations among the 20 folk concept scales for both males and females. Gough however states that

" because all of the scales assess facets of interpersonal functioning and because their intercorrelations are intended to mirror folk usage of the same concepts, non-zero and positive correlations among the scales are to be expected, and are in fact desirable " (Gough, 1985, p. 32).

Table Four

Relationships Between Old and New Scales of the CPL

scale name	No. of items: In old scale	Dropped	Added	In new scale	Corr. between old and new scales: males (n=1000)	females (n=1000)
Do	46	10	0	36	.97	.97
Cs	32	4	0	28	.98	.98
Sy	36	4	0	32	.99	.98
Sp	56	18	0	38	.95	.96
Sa	34	6	0	28	.93	.94
In*	-	-	-	30	-	-
Em*	-	-	-	38	-	-
Re	42	6	0	36	.98	.98
So	54	8	0	46	.98	.98
Sc	50	12	0	38	.98	.98
Gi	40	5	5	40	.97	.96
Cm	28	3	13	38	.92	.88
Wb	44	6	0	38	.98	.98
To	32	10	10	32	.91	.93
Ac	38	10	10	38	.93	.93
Ai	32	7	11	36	.92	.93
Ie	52	12	2	42	.97	.97
Py	22	0	6	28	.95	.96
Fx	22	0	6	28	.96	.96
F/M	38	6	0	32	.95	.94

*scale not included in previous edition of the CPL

(Gough 1987)

Reynolds and Nichols (1977) carried out factor analytic studies on the CPI and discovered that intercorrelations gave rise to two major factors, neuroticism - stability and extraversion - introversion. They suggested that this may be all that the scales are measuring. Further analysis found that predictions from the earlier version's scales did not add significantly to the prediction from the two factors. More recently factor analysis has been calculated on the 20 scales which clearly defined four factors; extraversion, control, flexibility, and consensuality (Gough, 1987)

-RELIABILITY AND VALIDITY-

Reliability data for the scales as given in the manual appear acceptable. For University students the lowest alpha coefficient was 0.52 for self-acceptance (Sa) and the highest was 0.85 for level of realisation (V3).

Hogan and Hogan (1989) suggest that the CPI is the most carefully validated measure of normal personality among the currently available commercial inventories. Many construct validation studies exist an the evidence indicates that the scales measure what their titles suggest.

-USERS-

The CPI is not constructed to predict behaviour to a high degree of accuracy in narrowly defined situations, but instead scales show moderate correlations to a wide range of criteria.

In a pamphlet produced by the Consulting Psychologists Press, California, it is claimed that the CPI be used for educational or vocational assessment and counselling, for industrial settings in areas such as selection, advising and counselling, for training settings, for clinical assessment and for cross-cultural and other research. Such research includes academic achievement at all educational levels; drop out in high school, college or graduate school; honesty and dependability at work; creativity; life span outcomes and self-realisation; vulnerability to coronary attack; good health in later life; and sleep-related phenomena.

-DESCRIPTION OF THE CPI-

The CPI was originally devised by Gough in 1956 and was revised in 1987. The revision was more of a fine tuning where the original 18 scales were retained and two new scales, independence and empathy were added. Of the original 480 items 18 were eliminated and 29 were modified to reflect current usage, simplify wording and reduce sexist or other bias. The revised handbook claims that all of the revised scales correlate .95 or above with the original scales. The scales are

as follows; Dominance (Do), Capacity for status (Cs), Sociability (Sy), Social Presence (Sp), Self-acceptance (Sa), Independence (In), Empathy (Em), Responsibility (Re), Socialisation (So), Self-control (Sc), Good Impression (Gi), Communality (Cm), Sense of Wellbeing (Wb), Tolerance (To), Achievement via Conformance (Ac), Achievement via Independence (Ai), Intellectual Efficiency (Ie), Psychological-mindedness (Py), Flexibility (Fx), and Femininity / Masculinity (F/M).

Three of the scales (Wb, Gi, and Cm) have been devised to detect protocols that are invalid because of faking or carelessness. The three scales are also included in interpreting the subject's personality.

Some structuring of the CPI is included in the 1987 revision. Three major vectors have been developed, which when used in combination give a structural model of personality. These vectors are; Internality (V1), Norm-favouring (V2), and Self-realisation (V3).

Combining V1 and V2 places individuals into one of four types, and V3 gives the degree to which the potentials associated with these four types has been realised by the individual.

-ADMINISTRATION OF THE CPI-

Administration typically requires 45 - 60 minutes although no time limits should be given. It is designed for ages 14 years to adult. The form of the CPI is forced choice format, where subjects get to answer either 'true' or 'false' to 462 questions. Subjects respond to the test items by putting their answer on a provided answer sheet. Once completed the administrator calculates the raw score for each of the 20 scales and the three vectors by the use of stencils. The raw scores for each scale are then plotted onto a profile form which shows their corresponding sten value. Ranges for each scale vary, but for all scales fall between zero and ten. Regardless of the highest and lowest score possible, for every scale a score of five is considered a balance between each end of the scale. The raw scores for the three vectors are plotted onto a graph and scale to discover the subject's type and level of self-realisation. For the present study the three vector scales were not included in the calculations because they do not appear in the main body of the CPI profile form

Interpretation is surmised via the study of the profile form or through computer generated reports available from the publisher. The interpretation guide describes each scale as well as the scale combinations.

2. The CPI Profile Form.

The profile form for the CPI (see appendix E) is incomprehendable in its present

form for the task of subjects self assessing their personalities. In its present form the profile form has only abbreviations, such as 'Do' for 'dominance', and no descriptions of what dominance means. Usually the profile form is used by the examiner to plot the subjects' results from the CPI. So that subjects could plot their scores as they see themselves the profile form was revised and abbreviations replaced with descriptions of what each scale means. For example 'Do' (Dominance) was replaced with 'unassuming, not forceful' on the lower end of the scale and 'confident, assertive, dominant, task-oriented' on the higher end of the scale (see appendix F).

On a number of factors it was not possible in the test to score above and below certain scores and so borders were made on the newly devised CPI profile form so subjects could not give themselves ratings impossible to get. These borders differed very slightly for males and females. For most scales the range of responses was between one and nine.

3. The General Questions Form (GQF) (see appendix B).

The GQF was designed by the experimenter to measure subjects previous experience at self assessment and their confidence at self assessing their personalities as well as finding out general information about the subjects such as age, gender and ethnic group.

4.03 ETHICAL CONSIDERATIONS

1. Informed consent.

Informed consent was obtained in two ways. Firstly subjects agreed to participate and secondly they signed a consent form (see appendix C).

2. Debriefing.

Subjects were debriefed once the experiment was completed. A written and verbal explanation was given and discussion was encouraged.

The procedure and ethical considerations were discussed and approved by the Massey University Ethics Committee.

4.04 PROCEDURE

Subjects placed themselves into one of two experimental conditions, depending on which time slots suited their timetables. The experiment had two parts to it spaced one week apart. There were also two conditions, condition A and condition B. In condition A subjects did the personality test first and the self assessment one week later. In condition B subjects did the self assessment first and the personality test one week later. This was done to test for any ordering effects.

CONDITION A.

Part One.

Subjects in condition A (16 males and 14 females) were administered the CPI in a classroom. They were given code numbers to use for identification instead of using their names for the sake of anonymity. They were also asked to place their date of birth on their answer sheets as a back up for identification in case subjects forgot their code number the following week when doing the second part of the experiment.

Before doing the personality test subjects filled in a consent form and were then told to read the instructions on the front page of the CPI carefully. They were reminded that they were free to drop out of the experiment at any time without giving any explanation.

The subjects could start when they were ready. The experimenter remained seated in the room with the subjects throughout the duration of the test, which took on average about 50 to 60 minutes to complete. On completion subjects were reminded that the second part to the experiment was a week away and that the experimenter would call to remind them the night before. They were also told to leave their names and experimental code numbers on a drawn up piece of paper on the experimenter's desk if they were interested in getting their individual results back.

Part two.

Part two was completed in the same classroom at the same time of day a week later. The subjects were given the GQF to complete. The subjects were then asked to rate themselves as they believe themselves to be on the revised CPI profile form.

CONDITION B.

The subjects in condition B were nine males and seven females. Conditions were identical to those in condition A except the subjects in condition B did part two, the GQF and self assessment first, and part one, the personality test a week later. For both condition A and B the consent form was completed before the experiments commenced. This meant that the subjects in condition A did this before the personality test and subjects in condition B did this before the GQF and their self assessment on the CPI profile form.

4.05 DEBRIEFING

Debriefing occurred immediately after the subjects had completed the second part of the experiment. The debriefing consisted of a verbal discussion and outline of the experimental aim and procedure and a written verification of these. The subjects were encouraged to discuss any points or ask any questions.

4.06 FEEDBACK

Feedback was given approximately three weeks later when subjects collected their individual results. Overall results had not been calculated at this stage and subjects were informed where and when they would be posted.

CHAPTER FIVE

RESULTS

Experiment Two

Subjects were generally confident at rating their personalities on the profile form. On a scale of one to ten (with ten being extremely confident), the mean and modal responses were seven and a half and eight respectively, with only six and a half percent of subjects rating their confidence at the task below five.

86 percent of subjects had had no previous experience at self assessment of this nature and therefore this variable was excluded from further analysis.

Twenty analyses of variance (ANOVAs) were calculated to test whether condition and gender had any influence on subjects' accuracy at rating their personalities as measured by the CPI. The dependent variables were the total of the differences between each subject's test score and self assessment score for each scale. Hence 'DIFFDo' equals Σ(TESTDo - SELFDo). Two tailed correlations were performed across these 'diff' values and significant correlations were found to be few. Therefore factorial ANOVAs were calculated rather than MANOVAs (multiple analysis of variance). Results were found to be non-significant ($p < 0.004$ ie .05/20 where 20 equals the number of ANOVAs calculated) indicating that gender and condition had no influence on subjects' accuracy at rating their personalities

as measured by the CPI.

The hypothesis that subjects are able to assess their personalities as measured by the CPI was tested by calculating two tailed correlations ($p < .01$) between the test scores and self assessment scores on each of the 20 scales. Of the 20 correlations six were found to be significant; Sy, Do, Fm, Sp, Sa, and Ac. If the alpha level was altered to .05 another four scales would be significant; Ie, In, Re and Cs. Because 20 correlations were calculated however if alpha equals .05 it is possible one scale would become significant by chance. Therefore the hypothesis that subjects are able to assess their personalities as measured by the CPI is partially supported.

Scatterplots were run between the CPI test scores and the subjects' self assessment scores for the 20 scales and means and standard deviations were calculated (see table five) to aid in interpretation of the correlations.

Sy (sociability) sociable, likes to be with people, friendly versus shy, feels uneasy in social situations, prefers to keep in the background.
The correlation between the CPI tests scores and the subjects' self assessment scores for Sy is 0.6613 ($p= 0.000$). The scatterplot clearly shows a positive linear relationship with much of the data very central and no outliers. The mean Sy test score is 5.57 and the mean self assessment score is 5.72.

Table Five

Correlations between the CPI test scores and corresponding self assessment scores and means and standard deviations for the CPI test scores, self assessment scores and scores of the difference between the two.

CPI scale	corr of test with self	p value	test score mean	test score SD (n=45)	self assess mean (n=45)	self assess SD	Diff scores	Diff SD
Sy	0.6613	0.000	5.57	0.83	5.72	1.14	-0.15	0.86
Do	0.6366	0.000	5.61	1.21	5.87	1.34	-0.26	1.09
Fm	0.5078	0.000	4.54	1.16	5.21	1.44	-0.67	1.31
Sp	0.4356	0.002	5.53	1.05	5.53	1.64	-.00	1.51
Sa	0.4089	0.005	5.62	0.88	5.17	1.20	0.45	1.16
Ac	0.3964	0.006	4.90	0.72	5.34	1.01	-0.43	0.98
Ie	0.3404	0.021	5.03	0.86	4.90	1.13	0.13	1.17
In	0.3133	0.034	5.32	0.77	5.65	1.08	-0.33	1.11
Re	0.3097	0.036	4.67	0.78	5.61	0.94	-0.94	1.02
Cs	0.3082	0.037	5.07	0.81	6.14	0.98	-1.07	1.06
So	0.2204	0.141	4.81	0.77	4.74	1.35	0.07	1.39
Wb	0.2060	0.170	5.23	0.66	4.84	0.97	0.39	1.05
Fx	0.1852	0.218	5.37	1.09	5.66	1.06	-0.30	1.37
Gi	0.1663	0.269	5.13	0.91	4.87	1.26	0.26	1.42
Sc	0.1624	0.281	4.82	0.85	5.08	1.27	-0.26	1.41
To	0.1066	0.481	5.13	0.81	5.52	0.99	-0.39	1.22
Em	-0.0546	0.718	5.63	0.79	5.99	0.94	-0.36	1.26
Ai	-0.0626	0.679	5.38	0.66	5.24	1.27	0.14	1.47
Cm	-0.1006	0.506	5.18	0.53	3.92	1.35	1.27	1.50
Py	-0.1695	0.260	5.10	0.63	5.64	1.31	-0.55	1.54

Do (dominance) confident, assertive, dominant, task-oriented versus unassuming, not forceful.

The CPI test scores and the self assessment scores correlate 0.6366 (p= 0.000) for this scale. The scatterplot indicates a positive linear relationship with much of the data central with no outliers. The mean CPI test scores and self assessment scores are 5.61 and 5.87 respectively.

Fm (femininity/masculinity) sympathetic, helpful, sensitive to criticism, tends to interpret events from a personal point of view, often feels vulnerable versus decisive, action-orientated, takes the initiative, not easily subdued, rather unsentimental.

The Fm scale yields the third strongest correlation between the CPI test scores and the subjects' self assessment scores of 0.5078 (p= 0.000). The scatterplot indicates a positive linear relationship. The mean CPI test scores and self assessment scores respectively are 4.54 and 5.21.

Sp (social presence) self assured, spontaneous, a good talker, not easily embarrassed versus cautious, hesitant to assert own views or opinions, not sarcastic or sharp tongued.

The correlation for scale Sp between the CPI test scores and the subjects' self assessment scores is 0.4356 (p= 0.005). The scatterplot indicates a positive linear relationship. Much of the data is central in the scatterplot with no outliers. The mean CPI test scores and self assessment scores are 5.23 and 5.53 respectively.

Sa (self-acceptance) has good opinion of self, sees self as talented, and as personally attractive versus self doubting, readily assumes blame when things go wrong, often thinks others are better.

The CPI test scores and self assessment scores correlate 0.4089 (p= 0.005) for this scale, with the group means respectively being 5.62 and 5.17. The scatterplot indicates a positive linear relationship. The range of the data is very small with much of the data clustered centrally. There are no outliers.

Ac (achievement via conformance) has strong drive to do well, likes to work in settings where tasks and expectations are clearly defined versus has difficulty in doing best work in situations with strict rules and expectations.

The Ac scale yields the next largest correlation between the CPI test scores and the subjects' self assessment scores of 0.3964 (p= 0.006). The scatterplot is similar to the Sa (self-acceptance) scatterplot in that the data range is very small and clustered with no outliers. The mean CPI test score is 4.90 and self assessment score is 5.34.

Ie (intellectual efficiency) efficient in use of intellectual abilities, can keep on at task where others might bored or discouraged versus has hard time getting started on things, and seeing them through to completion.

The correlation between the CPI tests scores and the self assessment scores for Ie is 0.3404 (p= 0.021). The scatterplot indicates a weak linear relationship. Again the data is central with a small range. The mean CPI test score is 5.03 and the mean self assessment score is 4.90.

In (independence) self-sufficient, resourceful, detached versus lacks self confidence, seeks support from others.

The CPI test scores and subjects' self assessment scores correlate 0.3133 (p= 0.034) for this scale. The scatterplot indicates a weak linear relationship with the data being central with a small range and no outliers. The means for the CPI test scores and subjects' self assessment scores are 5.32 and 5.65 respectively.

Re (responsibility) responsible, reasonable, takes duties seriously versus not overly concerned about duties and obligations, may be careless or lazy.

The correlation between the CPI test scores and subjects' self assessment scores for Re is 0.3097 (p= 0.036). Again the scatterplot indicates a weak linear relationship with the data remaining central and no outliers. The mean test score is 4.67 and self assessment score is 5.61.

Cs (capacity for status) ambitious, wants to be a success, independent versus unsure of self, dislikes direct competition.

The Cs scale yields a correlation of 0.3082 (p= 0.037) between the CPI test scores and the subjects' self assessment scores. The scatterplot does not clearly indicate a linear relationship. The data is central with a small range and no outliers. The means for the CPI test scores and subjects' self assessment scores are both 5.53.

So (socialisation) comfortably accepts ordinary rules and regulations, finds it easy to conform versus resists rules and regulations, finds it hard to conform, not conventional.

Wb (well-being) feels in good physical and emotional health, optimistic about the future versus concerned about health and personal problems, worried about the future.

Fx (flexibility) flexible, likes change and variety, easily bored by routine life and everyday experience, may be impatient, and even erratic versus not changeable, likes a steady pace and well organised life, may be stubborn and even rigid.

Gi (good impression) wants to make a good impression, tries to do what will please others versus insists on being him/her self, even if this causes friction or problems.

Sc (self-control) tries to control emotions and temper, takes pride in being self disciplined versus has strong feelings and emotions, and makes little attempt to hide them, speaks out when angry or annoyed.

To (tolerance) is tolerant of others' beliefs and values, even when different from or counter to own beliefs versus not tolerant of others, sceptical about what they say.

These six scales yielded non-significant correlations between the CPI test scores and the subjects' self assessment scores. Respectively they are 0.2240 (p= 0.141), 0.2060 (p= 0.170), 0.1852 (p= 0.218), 0.1663 (p= 0.269), 0.1624 (p= 0.281) and 0.1066 (p= 0.418). The scatterplots for these scales do not indicate linear relationships. The data range tends to be small with no outliers.

Em (empathy) comfortable with self and well accepted by others, understands the feelings of others versus ill at ease in many situations, unempathic.

Ai (achievement via independence) has strong drive to do well, likes to work in

settings that encourage freedom and individual initiative versus has difficulty doing best work in situations that are vague, poorly defined, and lacking in clear-cut methods and standards.

Cm (communality) fits in easily, sees self as quite an average person versus sees self as different from others, does not have the same ideas, preferences etc.

Py (psychological-mindedness) more interested in why people do what they do than in what they do, good judge of how people feel and what they think about things versus more interested in the practical and concrete than abstract, looks more at what people do than what they feel or think.

These four scales yielded negative non-significant correlations between the CPI test scores and the subjects' self assessment scores with Py being the worst. The correlations were Em -0.0546 (p= 0.718), Ai -0.0626 (p= 0.679), Cm -0.1006 (p= 0.506) and Py -0.1695 (p=0.260). The scatterplots for Cm and Py show a very small range of subjects' CPI test scores which hinders any possibility of a significant correlation. The scatterplot for Em shows a central cluster of data with a small range and no outliers, and suggests no linear relationship. The scatterplot for Ai is somewhat deceiving in that the relationship appears clear and positive. However there is no relationship between the subjects' self assessment scores and the CPI test scores for this scale. All of the data is very central with no outliers.

5.01 CONCLUSION

The hypothesis that gender and condition would not affect subjects' ability to assess their personalities as measured by the CPI was supported. However, the

hypothesis that subjects would be able to assess their personalities as measured by the CPI was only partially supported. Six of the twenty scales (Sy, Do, Fm, Sp, Sa and Ac) found this to be true when $p<0.01$. This applied also to scales Ie, In, Re and Cs when the alpha level was raised to 0.05. For the remaining ten scales no significant relationship between subjects' self assessment scores and CPI test scores was found. However, one clear trend that did emerge was the small range of data and the lack of outliers. On a scale ranging between zero and ten the majority of the test and self assessment scores for most factors remained between three and seven, therefore significantly hindering the size of any correlations. This small central cluster of data suggest too that even though the correlation may be small or non-significant subjects are able to assess their personalities similarly to the personality test. This is reflected through the mean and standard deviation values.

CHAPTER SIX

METHOD 3 (SOCIAL DESIRABILITY CHECK)

Experiment Three

6.01 INTRODUCTION

The aim of the third experiment was to find if certain responses are considered to be more socially desirable on the profile forms for both the CPI and 16PF. It was hypothesised that factors on the 16PF and scales on the CPI are not socially neutral. By this it is meant that one pole of each factor or scale is more socially desirable than the opposite pole. It was hypothesised in experiments one and two that the subjects' personality test scores and self assessment scores would not be significantly different. It is hypothesised in this experiment that the subjects will not rate themselves, when doing their self assessments, in a socially desirable way (ie towards the more socially attractive pole).

6.02 SUBJECTS

The subjects were 31 volunteers from a third year psychology class at Massey University. There were seven (23%) males and twenty four (77%) females. The age range was nineteen to forty two years with the mean and modal ages being twenty two and a half years and twenty one years respectively.

6.03 MATERIALS

1. 16PF Profile form (see appendix A).

2. New CPI Profile form (see appendix F).

6.04 PROCEDURE

The aim of the experiment was explained to the subjects. They were told that social desirability may have influenced how people in experiment one and two did their self assessments on the profile forms. The subjects were then asked to rate on the profile forms for the 16PF and CPI what they considered to be an 'ideal' person. The subjects took about five to ten minutes to do this. The subjects were asked not to identify who they were but to give their age and gender to compare this subject population to the subject population of experiment one and experiment two.

6.05 ETHICAL CONSIDERATIONS

Consent:

The subjects were asked if they would like to participate in the present experiment. They were informed that if they did not want to they did not have to.

6.06 FEEDBACK

The subjects were informed of where and when the results would be available.

CHAPTER SEVEN

RESULTS

Experiment Three

The hypothesis that items in the personality tests are not socially neutral has been supported by both the California Psychological Inventory (CPI) and the Sixteen Personality Factor Questionnaire (16PF) results. Tables six and seven show ratings of what is considered to be an 'ideal' person for each factor and scale and gives the minimum and maximum response given for that factor or scale. For the factor or scale to be socially neutral a rating of approximately 5.5 for the 16PF and 5.0 for the CPI was needed.

The hypothesis that subjects would not respond in a socially desirable way when doing their self assessment was not strongly supported by subjects who did the 16PF. Nine out of sixteen factors found significant differences ($p < 0.01$) between the means for the test scores on the 16PF and the corresponding self assessment scores. These factors are A, B, C, G, I, L, N, O, and Q4. All of these nine factors found that the self assessment score was closer to the ideal score than the corresponding personality test score. Table eight shows the personality test score means, self assessment means, and 'ideal' means for these nine factors. This indicates that on these nine factors the subjects rated themselves on the profile form when doing their self assessments in a socially desirable way.

Table Six

Means, standard deviations and the range of scores for what is considered to be an 'ideal' person on the 16PF.

16PF	'Ideal'	'Ideal'	Range	
Scale	mean	SD	Min	Max
A	7.710	1.24	5	10
B	8.323	1.49	5	10
C	7.484	1.41	5	10
E	5.600	1.58	1	8
F	7.258	1.00	5	10
G	7.129	1.23	4	9
H	7.484	1.21	6	10
I	4.613	1.23	2	7
L	3.742	1.77	1	7
M	5.419	1.77	2	9
N	4.323	1.70	1	7
0	2.452	0.96	1	4
Q1	7.710	1.37	5	10
Q2	7.613	1.23	5	10
Q3	6.387	1.59	4	10
Q4	2.806	1.08	1	5

Table Seven. Means, standard deviations and the range of scores for what is considered to be an 'ideal' person on the CPI.

CPI	'Ideal'	'Ideal'	Range	
Scale	mean	SD	Min	Max
Do	6.258	1.37	3	8
Cs	7.323	1.22	5	10
Sy	7.774	1.33	5	10
Sp	7.226	1.56	4	10
Sa	8.000	1.29	5	10
In	7.613	1.23	5	10
Em	8.806	1.20	6	10
Re	7.806	1.30	5	10
So	5.742	1.90	3	10
Sc	4.742	2.19	1	10
Gi	3.645	1.54	1	7
Cm	4.258	1.98	1	8
Wb	7.871	1.57	5	10
To	8.452	1.55	5	10
Ac	6.710	2.24	2	10
Ai	7.677	1.58	5	10
Ie	7.839	1.39	5	10
Py	7.452	1.73	4	10
Fx	6.452	1.59	3	10
F/M	5.065	1.41	2	7

Table Eight

The test, self assessment and 'ideal' means for the nine factors of the 16PF that suggest a social desirability bias.

16PF Scale	test mean (n=57)	self mean (n=57)	t-test p value	'ideal' mean (n=31)
A	5.1754	7.3509	0.00	7.710
B	6.0702	7.5789	0.00	8.323
C	5.4211	6.8070	0.00	7.484
G	4.1404	6.9649	0.00	7.129
I	5.7365	4.7719	0.01	4.613
L	6.3860	4.3158	0.00	3.742
N	3.5965	4.9825	0.00	4.323
O	5.842	4.5263	0.00	2.452
Q4	5.7368	4.6316	0.00	2.806

Table Nine

The test, self assessment and 'ideal' means for the five scales of the CPI that suggest a social desirability bias.

CPI Scale	test mean (n=46)	self mean (n=46)	t-test p value	'ideal' mean (n=31)
Cs	5.0674	6.1413	0.00	7.323
Re	4.6636	5.6087	0.00	7.806
Cm	5.1891	3.9239	0.00	4.258
Ac	4.9043	5.3370	0.01	6.710
Fm	4.5370	5.2065	0.00	5.065

The hypothesis that subjects would not respond in a socially desirable way when doing their self assessment was more strongly supported by subjects who did the CPI. On five of the twenty scales significant differences were found between the means for the personality test scores and the corresponding self assessment scores. These scales are Cs, Re, Cm, Ac and Fm. All of these five scales found that the self assessment score was closer to the 'ideal' score than the corresponding personality test score. Table nine shows the means for the personality test scores, self assessment scores and 'ideal' scores for these five scales. This indicates that on these five scales the subjects rated themselves on the profile form when doing their self assessments in a socially desirable way.

7.01 CONCLUSION

The hypothesis that items in the personality tests are not socially neutral has been supported by both the California Psychological Inventory (CPI) and the Sixteen Personality Factor Questionnaire (16PF). It appears for both personality tests that for the majority of factors and scales one pole is more socially attractive than the other.

The hypothesis that subjects would not respond in a socially desirable way when doing their self assessment was not strongly supported by subjects who did the 16PF. Nine out of sixteen factors found significant differences ($p < 0.01$) between the means for the test scores on the 16PF and the corresponding self assessment scores. For the CPI five out of twenty scales found significant differences ($p < 0.01$) between the means for the test scores on the CPI and the corresponding self assessment scores.

CHAPTER EIGHT

DISCUSSION

Three experiments were performed in the present study; the question of whether subjects could accurately assess their personalities as measured by a personality test was asked in experiments one and two with the Sixteen Factor Personality Questionnaire (16PF) and the California Psychological Inventory (CPI) used as the personality tests. The third experiment was performed to find whether the subjects in experiment one and two responded when rating their personalities in a socially desirable way. The results were not clear cut and raise a number of issues for discussion.

8.01 EXPERIMENTS ONE AND TWO

For both experiments one and two no moderator effects were found for gender or condition. Previous experience at self assessment was dropped from the experiment as a potential moderator due to insignificant numbers having had any previous experience. Very few subjects were from a minority ethnic group and therefore this variable was dropped from any further calculations. Confidence at doing the self assessment ratings on the profile forms was high for all subjects and therefore this variable too was excluded from further calculations.

For the 16PF the subjects were accurately able to assess their personalities as

measured by this personality test on five of the sixteen factors. For the CPI this was the case for six of the twenty scales. Four more scales on the CPI found significant correlations between the subjects' self assessment scores and their corresponding personality scores when the alpha level was raised to 0.05. Insignificant correlations between the subjects' self assessment scores and personality test scores on the CPI was strongly influenced by the small range of data and the lack of outliers; on a scale ranging between zero and ten nearly all of the test and self assessment scores for most factors remained between three and seven, therefore significantly hindering the size of any correlations. However, for many of these insignificant correlations the scatterplots between the subjects' self assessment scores and corresponding personality test scores appear to have a positive linear relationship.

8.02 FACTOR AND SCALE PATTERNS

Although the factors of the 16PF are different to the scales of the CPI definite patterns emerged indicating that certain personality dimensions are either easier to self assess or are less prone to bias. This was an 'unpredicted' discovery but one worth looking into.

The highest correlations between the subjects' self assessment scores and their corresponding personality test scores were *boldness* for the 16PF and *sociability* for the CPI. This factor and scale appear from their descriptions to be very similar;

Boldness (H): shy, restrained, timid, threat-sensitive versus venturesome, socially bold, uninhibited, spontaneous.

Sociability (Sy): sociable, likes to be with people, friendly versus shy, feels uneasy in social situations, prefers to keep in the background.

The second largest correlation between the subjects' self assessment scores and corresponding personality test score for the 16PF was *insecurity*. This factor is similar to the *social presence* scale on the CPI, which yielded the fourth largest correlation between the subjects' self assessment scores and CPI test scores. This factor and scale are defined as follows;

Insecurity (O): self-assured, confident, serene versus apprehensive, self-reproaching, worried, troubled.

Social Presence (Sp): self-assured, spontaneous, a good talker, not easily embarrassed versus cautious, hesitant to assert own views or opinions, not sarcastic or sharp tongued.

The second largest correlation for the CPI and the fifth largest correlation for the 16PF between the subjects' self assessment and personality test scores were the *dominance* factor and scale. These were defined as follows;

Dominance (E): humble, mild, accommodating, conforming versus assertive,

aggressive, stubborn, competitive.

Dominance (Do): confident, assertive, dominant, task oriented versus unassuming, not forceful.

The third largest correlation for the 16PF was *conformity* which is similar to the *achievement via conformance* scale on the CPI, which yielded the sixth largest correlation from the CPI data. This factor and scale are defined as follows;

Conformity (G): expedient, disregards rules, feels few obligations versus conscientious, persevering, staid, moralistic.

Achievement via Conformance (Ac): has strong drive to do well, likes to work in settings where tasks and expectations are clearly defined versus has difficulty in doing best work in situations with strict rules and expectations.

All of these correlations reached a level of significance, and hence four out of five of the 16PF's significant correlations resembled four out of six of the CPI's significant correlations.

Adcock (1974) found the most promising factors of the 16PF to be *sensitivity (I), boldness (H), impulsivity (F)* and *conformity (G)*. Of these four factors three (H, F and G) were found in the present experiment to have significant correlations between the subjects' self assessment and personality test scores. Furthermore

Adcock (1974) found five factors (*shrewdness (N), self-sufficiency (Q2), radicalism (Q1), suspiciousness (L)* and *imagination (M)*) failed to achieve better than chance discrimination of items in each factor. The five lowest correlations found between the subjects' self assessment and personality test scores for the present experiment included four of these five factors; M, Q2, L and M.

These findings suggest that the correlations which did not reach levels of significance may not be due to subjects responding in a socially desirable way, but instead may be due to factors not measuring what they claim to. Hence as subjects did their self assessments their interpretation of the factors on the profile form did not coincide with the meanings attributed to these factors by the 16PF.

8.03 EXPERIMENT THREE

The aim of the third experiment was to see whether each end of each factor on the 16PF and each scale on the CPI is as socially desirable as the opposite end and then to find whether the subjects, when doing their self assessments, would rate themselves differently to their personality test scores and in a more socially desirable way. The results indicated for both the CPI and 16PF that for almost all scales and factors one end of the scale or factor was more socially desirable than the other. For the 16PF it was found that subjects rated themselves significantly differently to their corresponding personality test scores and, in fact, their self assessment scores reflected a socially desirable response for eight of the sixteen factors. For the CPI this occurred for only three of the twenty scales.

These results suggest response bias is prevalent although other explanations are likely. Because the 16PF found a much stronger bias than the CPI it could be due to a fault with the 16PF. Possibly the factors do not measure what they claim to measure (as suggested above) and/or subjects may have interpreted the descriptions of the factors on the profile form incorrectly. Kagan (1988) suggested that definitions in personality testing can often be misinterpreted depending on who is interpreting and under what conditions. The descriptions of the scales on the profile form for the CPI were much simpler to understand. Furthermore, the subjects when doing their self assessments on the profile forms for the 16PF often claimed difficulty in doing the self assessment because each end of each factor was not necessarily an opposite to the other end and at times the subjects felt both ends applied to some degree on some of the descriptions to them. This complaint did not arise with the CPI. This problem with the 16PF is due to the development of the 16PF being via factor analysis. Each factor has a weighting of a number of personality dimensions within that factor, and it is therefore possible to feel that descriptions on each end of the factor are not opposite and that a mixture of descriptions from both ends may apply to a respondent.

8.04 THE 16PF VERSUS THE CPI

An interesting finding from these three experiments has been the differences the two personality tests, the 16PF and CPI have had on the results. The most obvious reasons for this are the ones given above; less clear descriptions on the

16PF and different techniques of test development (factor analysis versus interpretational). Bull (1974) has stated that factor analysis may indicate structure but generally has nothing to do with the validity of the test in an applied setting.

The stability of the factor structure for the 16PF is questionable, with little replication for the factors in a number of settings having been demonstrated (Zuckerman 1985, Peterson 1965, Eysenck, White and Souief 1969, Eysenck 1969, and Howarth and Brown 1971). Furthermore the items per factor is very low (10 to 13 items per factor) and Adcock (1974) found many of these items were measuring related factors and not those for which they were designed. The CPI has 28 to 46 items per scale and evidence indicates the scales measure what their titles suggest.

Another problem with the 16PF is that its limits of applicability are not clearly defined. It was developed as a research instrument for assessing source traits but has been used in a wide range of situations. The CPI however was developed for use in a wide range of situations, but was not constructed to predict behaviour to a high degree of accuracy. Both tests were intended for use on a 'normative' rather than clinical population. However the interpretation manual for the 16PF makes a number of clinical statements.

Methodologically, a problem arose which may have caused bias in the results. The CPI scale range was 0-10, but on most factors extremely high or low scores were not possible to get. The subjects doing their self assessments on the CPI

profile form were therefore instructed as to what the actual range for each scale was. This cut down the range of possible responses. For the self assessment on the 16PF the range was consistently 1-10. Statistically this meant that correlations between self assessment and personality test scores for the CPI would be hindered due to a lack of outliers and a smaller range. However the scatterplots for the CPI data of self assessment versus personality test scores suggest a positive linear relationship for many of the scales, even though not all of these scales produced significant correlations.

8.05 FACTORS AFFECTING SELF ASSESSMENT.

A number of suggestions have been made to enhance accuracy of subjects' responses when doing self assessments. However most of these suggestions are applicable only to applied settings such as personnel selection, and are intended for ability testing as well as personality testing. For example telling people that their self assessments will be validated against other information and warning them about consequences of distortion was not viable in the present experiments.

Response validity scales as recommended by Hough, Eaton, Dunnette, Kamp and McCloy (1990) are a feature of the CPI, but not the 16PF for the form chosen (form A). No subjects who did the CPI were found to have distorted their answers on these scales, which suggests distortion on the personality tests is not a problem in this setting.

Encouraging motivation in a lab setting differs somewhat from encouraging motivation in an applied setting. The subjects were encouraged to answer honestly and were motivated by their interest in their personality test results, and their ability to predict their personalities as measured by a personality test. In an applied setting such as personnel selection applicants' motivation becomes questionable, especially at a time when jobs are difficult to obtain. Schrader and Osburn's (1977) suggestion to evaluate and encourage motivation therefore becomes more crucial. By involving applicants in the decision making of the selection procedure may be one way of facilitating this.

Very few of the subjects in the present experiments had any previous experience at self assessment. Whether this is due to the subjects being students, with many not having gone through a personnel selection procedure, or whether self assessment is uncommon is uncertain. It seems likely that suggestions made to facilitate recall of previous experience (Schrader and Osburn, 1977; Mabe and West, 1982) apply more to testing abilities rather than personalities.

8.06 IMPLICATIONS.

The most important question faced after calculating the results for the present experiment was which is a better and more accurate measure of a person's personality, their personality test results or their own self assessment of their personality as plotted on a profile form ? The strongest correlations between the subjects' self assessment and personality test scores are 0.65 for factor H on the

16PF and 0.66 for scale Sy on the CPI. The percentage of variance explained for these correlations in one variable (the personality test score) by another (the self assessment score) are 42 percent for factor H and 44 percent for scale Sy. The percentage of variance explained in one variable by another is, for these higher correlations, moderate. However, the percentage of variance explained in one variable by another for the majority of factors and scales is poor. These two measures of personality, the personality test scores and corresponding self assessment scores, are either measuring two different aspects of the same construct, or one of these measures may be measuring the construct poorly while the other is measuring it accurately.

It will be argued by users of personality tests that these tests are a more accurate measure of one's personality. Not only are they less susceptible to distortion and response bias (due to the use of subtle items), but they also take into account subjects' responses to a number of situations, and have a scientific base (ie have been developed through extensive research).

Research however has found the use of subtle items to be often less valid than obvious items (Wiener 1948, McCall 1958, and Duff 1965). To ask people to self assess themselves directly on to a profile form excludes any use of subtle items and, in fact, the scales and factors are interpreted entirely in an open fashion.

Secondly, personality tests may take into account a number of situations but, by the same token, there are a number of situations they exclude. To ask a direct

question once rather than touch around the topic with a number of possibly relevant questions could yield a more accurate response. This argument is strengthened by the fact that it is accepted within the literature on self that people are able to know and assess themselves (Bandura, 1978; Bem, 1972; Mischel, 1972; Wicklund, 1975).

Thirdly, the argument that personality tests have a scientific base and are therefore superior to less scientific methods, such as self assessment, may hold weight in some situations. However, the tests currently used for personnel selection were not developed to perform that job alone and are too general. For a personality test to be appropriately used for personnel selection the questions in the test need to be related to the particular job being selected for. More research is needed in this area of work-related personality tests.

Both measures of one's personality (the personality test scores and the self assessment scores on the profile form) are a form of self assessment, with the main differences being that personality tests use a less direct approach and a lot more questions (and also time) to retrieve what should be the same information. The finding that these two means of assessing one's personality yield different results leads one to question the validity of each method. It has been shown in this experiment and others (Hough et al, 1990) that social desirability does not strongly influence the subjects' responses. It seems therefore plausible that the more direct approach of getting people to self assess their personalities on a profile form may be no less accurate and certainly more straight forward than

giving them the personality test to do. It hardly seems appropriate to tell subjects " sorry, but you got your personality wrong" after comparing self assessment and personality test scores.

The question of which is better, the person's self assessment or the personality test score, so far is unanswered. To cast doubt on the use of personality tests, especially in settings such as personnel selection may cause an outcry from professionals that choose to use them. Studies in the past (Blinkhorn and Johnson, 1990; George and Smith 1986) that have ridiculed the use of such personality tests for personnel selection have met with some criticism by other professionals (Smith, 1993). It may be that a person's personality is an important predictor of a person's ability to perform certain tasks but only if the person's personality is being accurately measured.

8.07 FUTURE RESEARCH.

One of the objectives of this research was to explore further the application of self assessment when used as a selection tool. Although findings from the present experiments are not from an actual selection setting, they were intended as a starting point for future research. It is interesting to note that one personality test, the CPI, yielded more promising results. A number of explanations have been offered for this finding, and these need to be further explored.

The finding of certain traits common to both the CPI and 16PF yielding bigger

correlations between subjects personality test scores and self assessment scores needs further attention. It may be that these scales and factors are actually tapping into what they claim to be and the other scales and factors are not. Again, more research is needed in this area.

To replicate this study in a selection setting may help interpret the present results, clear up some of the issues raised by the present experiment, and cast new light on this relatively new area of using self assessment as a tool for personnel selection.

Comparing people's profile forms to their job performance would further investigate the validity of self assessment as a predictive tool for personnel selection.

8.08 CONCLUSION

The present study focused on the use of self assessment on personality tests and the implications for personnel selection. Although the hypotheses were only partially supported, an unexpected finding that similar scales and factors on the CPI and 16PF yielded higher correlations between the personality test scores and self assessment scores was found.

It is suggested from the findings that people do have difficulty at assessing their personalities as measured by personality tests. However, it appears that they do

not bias their responses by answering in a socially desirable way. This suggests that people do respond honestly when doing self assessments and therefore it is possible that the factors and scales that did not yield significant correlations between personality test scores and self assessment scores was because people's self assessments were a more accurate rating of their personalities than the personality test scores were.

REFERENCES

Adcock, N.V.(1974). Testing the test: How adequate is the 16PF with a New Zealand student sample? New Zealand Psychologist, 4, 2-10.

Adcock, N.V., Adcock, C.J., & Walkey, F.H. (1971).Basic Dimensions of Personality. I.A.A.P. Congress, Liege.

Allport, G.W. (1961). Pattern Growth In Personality. New York: Holt, Rinehart and Winston.

Allport, G.W., & Odbert, H.S.(1936). Trait names: A psycho-lexical study. Psychological Monographs, 47, 1-171.

Anastasi, A. (1982). Psychological Testing. (5th Ed.). New York: MacMillan.

Anderson, C.D., Warner, J.L., & Spencer, C.C.(1984). Inflationbiases in self-assessment examinations: Implications for valid employee selection. Journal of Applied Psychology, 69, 574-580.

Anon, (1987). The Sales Performance Indicator. Assessment and Development Services Ltd: Christchurch, New Zealand.

Ash, R.A.(1980). Self-assessments of five types of typing ability. Personnel Psychology, 33, 273-281.

Asher, J.J., & Sciarrino, J.A.(1974). Realistic Work Samples: A review. Personnel Psychology, 27, 519-534.

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.(1982b). The self and mechanisms of agency. In J.Suls (Ed.), Psychological Perspectives on the Self (Vol.1, p 3-39). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Baucom, D.H. (1985). Review of the California Psychological Inventory. In J.V.Mitchell,Jn.(Ed.), The nineth mental measurements yearbook (vol.1). Lincoln, NE: Buros Institute of Mental Measurements. (BRS Document Reproduction Service No. AN 0910-679).

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.

Blinkhorn, S., & Johnson, C.(1990). The insignificance of personality testing. Nature, 348, 671-672.

Blum, M.L.,& Naylor, J.C. (1968) Industrial Psychology. Its Theoretical and Social Foundations. New York: Harper and Row.

Bull, P.E.(1974). Should the 16PF be used in personnel selection? New Zealand Psychologist, 4, 11-15.

Butcher, J.N.(1985). Review of the Sixteen Personality Factor Questionnaire. In J.V.Mitchell,Jn.(Ed.), The nineth mental measurements yearbook (vol.1). Lincoln, NE: Buros Institute of Mental Measurements. (BRS Document Reproduction Service No. AN 0910-679).

Cattell, R.B. (1972). Real base true zero factor analysis. Multivariate Behavioural Research Monographs, 72, 1-162.

Cattell, R.B., Eber, H.W., & Tatsuoka, M.M.(1970). Handbook for the 16PF. Institute for Personality and Ability Testing. U.S.A.

Cattell, R.B., Eber, H.W., & Tatsuoka, M.M.(1980). Handbook for the 16PF. Champaign, Illinois: IPAT.

Dakin, S., & Armstrong, J.S.,(1989). Predicting Job Performance: A comparison of expert opinion and research finding. International Journal of Forecasting, 5, 187-194.

Day, D.V., & Silverman, S.B.(1989). Personality and job performance: Evidence of incremental validity. Personnel Psychology, 42, 25-36.

De Nisi, A.S., & Shaw, J.B.(1977). Investigation of the uses of self reports of abilities. Journal of Applied Psychology, 62, 641-644.

Duff, F.L.(1965). Item subtlety in personality inventory scales. Journal of Consulting Psychology, 29, 565-570.

Edwards, A.L.(1957). The Social Desirability Variable in Personality Assessment and Research. New York: Dryden.

Elliot, A.G.P.(1976). Fakers: A study of managers' responses on a personality test. Personnel Review, 5, 33-37.

Elliot, A.G.P.(1981). Some implications of lie-scale scores in real life selection. Journal of Occupational Psychology, 54, 9-16.

Eysenck, H.J. (1985). Review of the California Psychological Inventory. In J.V.Mitchell,Jn.(Ed.), The nineth mental measurements yearbook (vol.1). Lincoln, NE: Buros Institute of Mental Measurements. (BRS Document Reproduction Service No. AN 0910-679).

Eysenck, H.J., White, P.O., & Soueif, M.I.(1969). Factors in the Cattell Personality Inventory. In H.J.Eysenck & S.B.G.Eysenck (Eds.) Personality Structure and Measurement. San Diego, CA: Robert R. Knapp Publisher.

Fazio, R.H., & Zanna, M.P.(1978). Attitudinal qualities relating to the strength of the attitude-behaviour relationship. Journal of Experimental Social Psychology, 14, 398-408.

Fox, S., & Dinur, Y.(1988). Validity of self-assessment: A field evaluation. Personnel Psychology, 41, 581-592.

George, D.I.(1988). The Evaluation of Self-Assessment for Personnel Selection in New Zealand. PhD Thesis. Massey University Library, New Zealand.

George, D., & Smith, M.(1990). An empirical comparison of self assessment and organisational assessment in personnel selection. Public Personnel Management, 19(2), 175-190.

Ghiselli, E.E., & Barthol, R.P.(1953). The validity of personality inventories in the selection of employees. Journal of Applied Psychology, 38, 18-20.

Gough, H.G. (1968). An interpreter's syllabus for the CPI. In P.McReynolds (Ed.). Advances in Psychological Assessment (Vol 1). Palo Alto, CA: Science and Behaviour books.

Gough, H.G. (1985). A work orientation scale for the CPI. Journal of Applied Psychology, 70, 505-513.

Gough, H.G. (1987). California Psychological Inventory Administrators' Guide. Palo Alto, CA: Consulting Psychologists Press.

Guilford, J.P. (1959). Personality. New York: McGraw-Hill.

Guilford, J.P., & Zimmerman, W.S. (1982). Fourteen dimensions of temperament. Psychological Monographs,70, (Serial No.10).

Guion, R.M., & Gottier, R.F.(1965). Validity of personality measures in personnel selection. Personnel Psychology, 18, 135-164.

Hathaway, S.R., & McKinley, J.C.(1967). Minnesota Multiphasic Personality Inventory Manual for Administration and Scoring. New York: Psychological Corporation.

Hesketh, B.(1974). Survey on Testing In Industry. Wellington, New Zealand: New Zealand Institute of Personnel Management.

Hogan, J., and Hogan, R. (1989). How to measure employee reliability. Journal of Applied Psychology, 74(2), 273-279.

Hough, L.M.(1989). Development of personality measures to supplement selection decisions. Advances in Industrial Organisational Psychology. B.J.Fallon, H.P.Pfister, J.Brebner (Eds). Elsevier Science Publishers B.V. North Holland.

Hough, L.M., Eaton, N.K., Dunnette, M.D., Kamp, J.D., & McCloy, R.A.(1990). Criterion-related variables of personality constructs and the effect of response distortion on those validities. Journal of Applied Psychology, 1975, 581-595.

Howarth, E., & Brown, J.A. (1971). An item factor-analysis of the 16PF. Personality: An International Journal, 2(2), 119-139.

Hunter, J.E., & Hirsh, H.R.H. (1987). Applications of Meta-analysis. International Review of Industrial and Organisational Psychology. New York: Wiley.

Hunter, J.E., & Hunter, R.F.(1984). Validity and utility of alternative predictors of job performance. Psychological Bulletin, 96, 72-98.

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

Inkson, K.(1987). Organisational Behaviour: A review of New Zealand Research. New Zealand Psychologist, 16, 9-27.

Kagan, J. (1988). The meaning of personality predicates. American Psychologist, 43(8), 614-620.

Kriedt, P.H., & Dawson, R.I.(1961). Response set and the prediction of clerical job performance. Journal of Applied Psychology, 45, 175-178.

Landy, F.J.(1989) Psychology of Work Behaviour. (4th Ed.).Brooks/Cole, California.

Levine, E.L., Flory, A., & Ash, R.A.(1977). Self assessment in personnel selection. Journal of Applied Psychology, 62, 428-435.

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.

Maddi, S.R. (1972). Personality Theories: A Comparative Analysis. The Dorsey Press; Illinois.

Markus, H., & Wurf, E.(1987). The dynamic self concept. Annual Review of Psychology, 38, 299-337.

McCall, R.J.(1958). Face validity in the D scale of the MMPI. Journal of Clinical Psychology, 14, 77-80.

McCormick, E,J. & Tiffin, J.(1974). Industrial Psychology (6th Ed.). New Jersey: Prentice-Hall.

Mergaree, E.I.(1972). The California Psychological Inventory Handbook. San Francisco: Jossey-Bass.

Merrideth, G.M.(1968). Stereotypic desirability profiles for the 16Pf Questionnaire. Psychological Reports, 23, 1173-1174.

Michaelis, W., & Eysenck, H.I.(1971). The determination of personality inventory factor patterns and intercorrelations by changes in real life motivation. Journal of Genetic Psychology, 118, 223-234.

Mischel, W.(1972). Direct versus indirect personality assessment: Evidence and implications. Journal of Consulting and Clinical Psychology, 38, 319-324.

Mischel, W.(1976). Introduction to Personality. Holt, Rinehart and Winston, New York.

Ng, E.(1990). A Meta-analysis on the validity of personality tests used for personnel selection. M.A. Massey University Library, New Zealand.

Ng, E., & Smith, M.C.(1991). An updated meta-analysis on the validity of personality tests for personnel selection for personnel selection. Submitted for publication and under review.

Pearlman, K. (1985). Validity Generalisation: From theory to application. Paper presented at Centre for Human Resources Programs. Institute of Industrial Relations, University of California - Berkley.

Pearlman, K., Schmidt, F.L., & Hunter, J.E.(1980). Validity generalisation results for tests used to predict job proficiency and training success in clerical employees. Journal of Applied Psychology, 65, 373-406.

Peterson, D.R. (1965). Scope and generality of verbally defined personality factors. Psychological Review, 72, 48-59.

Primoff, E.S.(1980). The use of self-assessments in examining. Personnel Psychology, 33, 283-290.

Regan, D.T., & Fazio, R.H.(1977). On the consistency between attitudes and behaviour: Look at the method of attitude formation. Journal of Experimental Social Psychology, 13, 28-45.

Reilly, R.R., & Chao, G.T.(1982). Validity and fairness of some alternative employee selection procedures. Personnel Psychology,35, 1-62.

Reynolds, C.H., & Nicholls, R.C. (1977). Factor rates for the CPI: Do they capture the valid variance? Educational and Psychological Measurement, 37, 907-915.

Ruch, F.L., & Ruch, W.W.(1967). The K factor as a (validity) suppressor variable in predicting success in selling. Journal of Applied Psychology, 51, 201-204

Ryan, A.M., & Sackett, P.R. (1987). A survey of individual assessment practices by I/O psychologists. Personnel Psychology, 40, 455-488.

Schmitt, N., Gooding, R.Z., Noe, R.A., & Kirsch, M.(1984). Meta-analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. Personnel Psychology, 37, 407-422.

Schmitt, N., & Robertson, I.(1986). Personnel selection. Annual Review of Psychology, 41, 289-319.

Schrader, A.D., & Osburn, H.G. (1977). Biodata faking: Effects of induced subtlety and position specificity. Personnel Psychology, 30, 395-404.

Scott, W.A., % Johnson, R.C.(1972). Comparative validities of direct and indirect personality tests. Journal of Consulting and Clinical Psychology, 38, 301-318.

Smith, G.A., Dowling, P.J., & Barry, B.(1985). Australian managers' norms on the 16PF. Australian Psychologist, 20, 293-301.

Smith, M.C (personal communication, February 19, 1993)

Thornton, G.C.III.(1980). Psychometric properties of self-appraisals of job performance. Personnel Psychology, 33, 263-271.

Thornton, G.C.III, & Gierasch, P.F.III.(1980). Fakability of an empirically derived selection instrument. Journal of Personality Assessment, 44, 48-51.

Tyler, B., & Miller, K.(1986). The use of tests by psychologists: Report on the survey by BPS members. Bulletin of the British Psychological Society, 39, 405-410.

Wicker, A.W.(1969). Attitudes versus actions: The relationship of verbal and overt behavioural responses to attitude objects. Journal of Social Issues, 25, 41-78.

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, p 67-92). Hillsdale, New Jersey: Lawrence Erlbaum.

Wiener, D.N.(1948). Subtle and obvious keys for the Minnesota Multiphaisic Personality Inventory. Journal of Consulting Psychology, 12, 164-170.

Zalinski, J.S., & Abrahams, N.M.(1979). The effects of item context in faking personnel selection inventories. Personnel Psychology, 32, 161-166.

Zanna, M.P., Olson, J.M., & Fazio, R.H.(1980). Attitude behaviour consistency: An individual difference perspective. Journal of Personality and Social Psychology, 38, 432-440.

Zuckerman, M.(1985). Review of the Sixteen Personality Factor Questionnaire.In J.V.Mitchell,Jn.(Ed.), The nineth mental measurements yearbook (vol.1). Lincoln, NE: Buros Institute of Mental Measurements. (BRS Document Reproduction Service No. AN 0910-679).

APPENDICES

A. Sixteen Personality Factor Questionnaire (16PF) Profile Form.

B. General Questions Form (GQF).

C. Consent Form.

D.1 - D.16. The Scatterplots of Self Assessment scores versus Personality Test scores for the Sixteen Factors of the 16PF.

E. The California Psychological Inventory (CPI) Profile Form.

F. The Revised Profile Form for the CPI.

G.1 - G.16. The Scatterplots of Self Assessment scores versus Personality Test scores for the Twenty Scales of the CPI.

APPENDIX A

SIXTEEN PERSONALITY FACTOR QUESTIONNAIRE (16PF)

PROFILE FORM.

IPAT

16 P.F. TEST PROFILE

Name:

Comments:

FACTOR	Raw Score Form C	Raw Score Form D	Raw Score Total	Standard Score	LOW SCORE DESCRIPTION	HIGH SCORE DESCRIPTION
A					RESERVED, DETACHED, CRITICAL, ALOOF (Sizothymia)	OUTGOING, WARMHEARTED, EASY-GOING, PARTICIPATING (Affectothymia, formerly cyclothymia)
B					LESS INTELLIGENT, CONCRETE-THINKING (Lower scholastic mental capacity)	MORE INTELLIGENT, ABSTRACT-THINKING, BRIGHT (Higher scholastic mental capacity)
C					AFFECTED BY FEELINGS, EMOTIONALLY LESS STABLE, EASILY UPSET (Lower ego strength)	EMOTIONALLY STABLE, FACES REALITY, CALM, MATURE (Higher ego strength)
E					HUMBLE, MILD, ACCOMMODATING, CONFORMING (Submissiveness)	ASSERTIVE, AGGRESSIVE, STUBBORN, COMPETITIVE (Dominance)
F					SOBER, PRUDENT, SERIOUS, TACITURN (Desurgency)	HAPPY-GO-LUCKY, IMPULSIVELY LIVELY, GAY, ENTHUSIASTIC (Surgency)
G					EXPEDIENT, DISREGARDS RULES, FEELS FEW OBLIGATIONS (Weaker superego strength)	CONSCIENTIOUS, PERSEVERING, STAID, MORALISTIC (Stronger superego strength)
H					SHY, RESTRAINED, TIMID, THREAT-SENSITIVE (Threctia)	VENTURESOME, SOCIALLY BOLD, UNINHIBITED, SPONTANEOUS (Parmia)
I					TOUGH-MINDED, SELF-RELIANT, REALISTIC, NO-NONSENSE (Harria)	TENDER-MINDED, CLINGING, OVER-PROTECTED, SENSITIVE (Premsia)
L					TRUSTING, ADAPTABLE, FREE OF JEALOUSY, EASY TO GET ALONG WITH (Alaxia)	SUSPICIOUS, SELF-OPINIONATED, HARD TO FOOL (Protension)
M					PRACTICAL, CAREFUL, CONVENTIONAL, REGULATED BY EXTERNAL REALITIES, PROPER (Praxernia)	IMAGINATIVE, WRAPPED UP IN INNER URGENCIES, CARELESS OF PRACTICAL MATTERS, BOHEMIAN (Autia)
N					FORTHRIGHT, NATURAL, ARTLESS, UNPRETENTIOUS (Artlessness)	SHREWD, CALCULATING, WORLDLY, PENETRATING (Shrewdness)
O					SELF-ASSURED, CONFIDENT, SERENE (Untroubled adequacy)	APPREHENSIVE, SELF-REPROACHING, WORRYING, TROUBLED (Guilt proneness)
Q_1					CONSERVATIVE, RESPECTING ESTABLISHED IDEAS, TOLERANT OF TRADITIONAL DIFFICULTIES (Conservatism)	EXPERIMENTING, LIBERAL, ANALYTICAL, FREE-THINKING (Radicalism)
Q_2					GROUP-DEPENDENT, A "JOINER" AND SOUND FOLLOWER (Group adherence)	SELF-SUFFICIENT, PREFERS OWN DECISIONS, RESOURCEFUL (Self-sufficiency)
Q_3					UNDISCIPLINED SELF-CONFLICT, FOLLOWS OWN URGES, CARELESS OF PROTOCOL (Low integration)	CONTROLLED, SOCIALLY PRECISE, FOLLOWING SELF-IMAGE (High self-concept control)
Q_4					RELAXED, TRANQUIL, UNFRUSTRATED (Low ergic tension)	TENSE, FRUSTRATED, DRIVEN, OVERWROUGHT (High ergic tension)

STANDARD TEN SCORE (STEN)

Average

1 2 3 4 5 6 7 8 9 10

A sten of 1 2 3 4 5 6 7 8 9 10 is obtained by about 4% 7% 12% 17% 20% 20% 17% 12% 7% 4% of adults

APPENDIX B

GENERAL QUESTIONS FORM.

Have you had any previous experience at self-assessment? If yes please specify.

Name:

What is your date of birth? / /

Please indicate your sex by circling the appropriate MALE / FEMALE.

To which ethnic group do you consider you belong? (Please circle one of the following)

N.Z European Maori Asian Polynesian

Other (please state).....

On a scale of 1 to 10 how confident are you at doing this self assessment? (where 1= not at all and 10 = very confident).

1 2 3 4 5 6 7 8 9 10

APPENDIX C

CONSENT FORM.

I ..... consent to being a subject in an
(PRINT NAME)

experiment on personality tests. I understand that I may

drop out of the experiment at any time I wish and need not

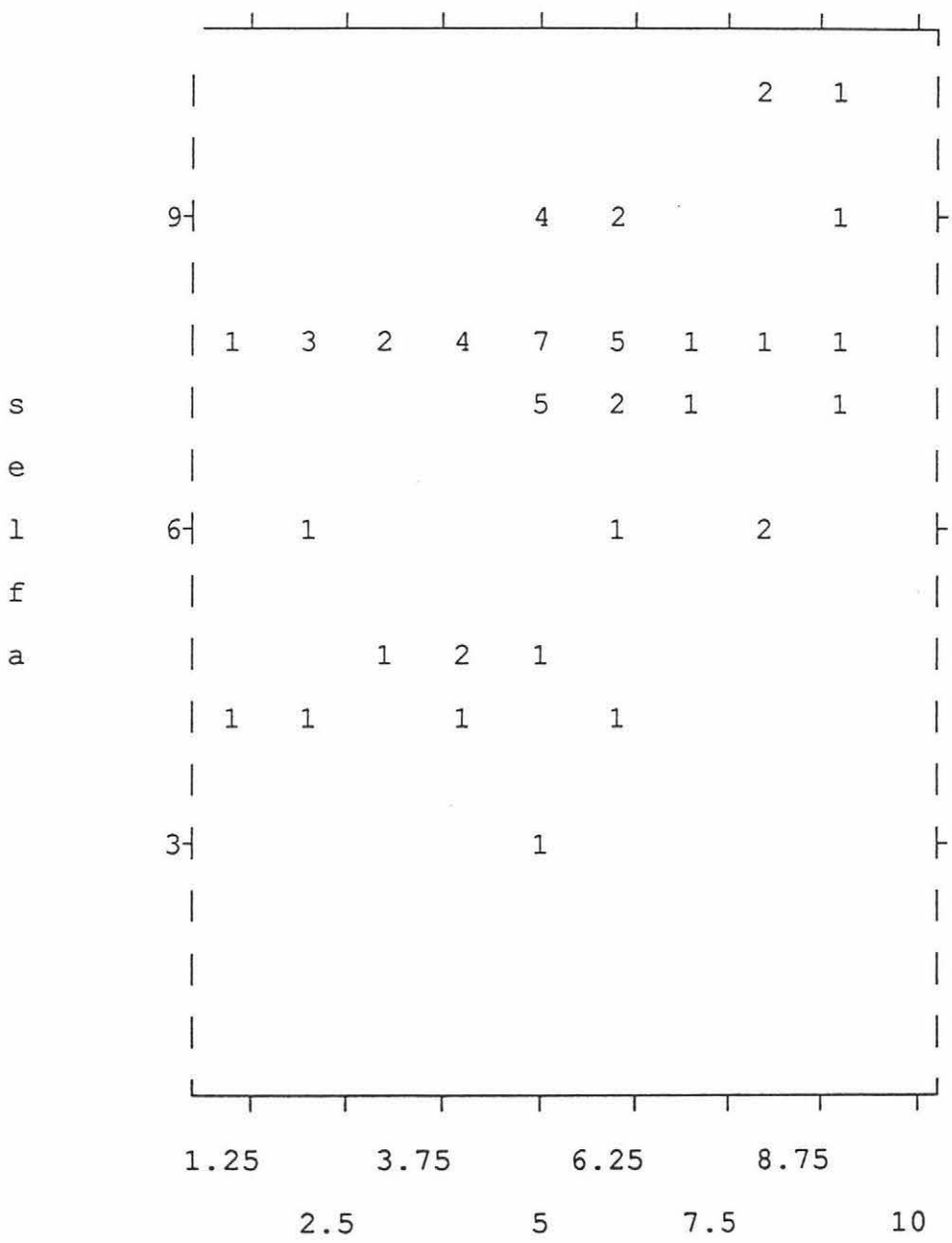
give any explanation.

Signed .....

Date .....

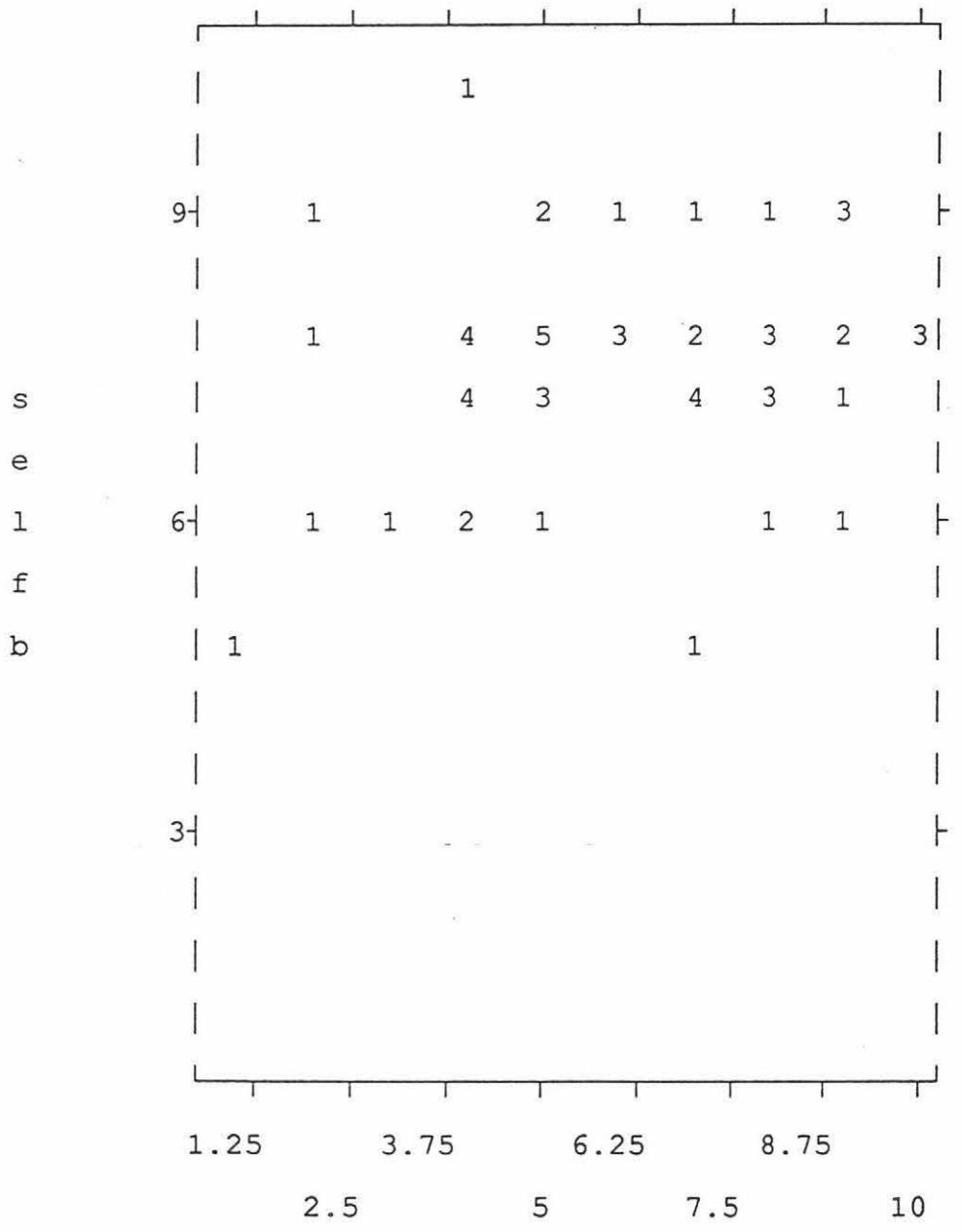
APPENDIX D.1

Scatterplot of self assessment scores and personality test scores for factor A on the 16PF.



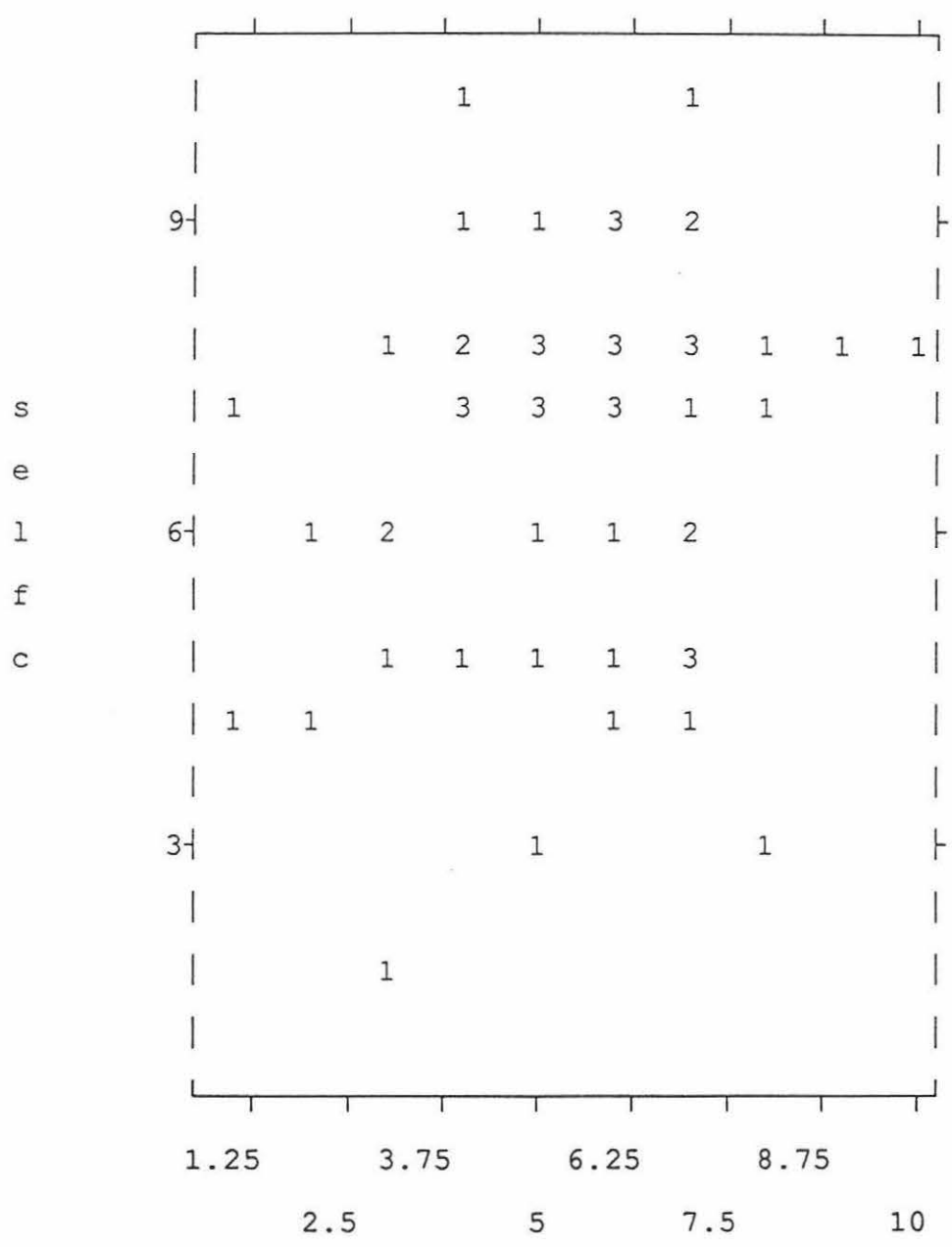
APPENDIX D.2

Scatterplot of self assessment scores and personality test scores for factor B on the 16PF.



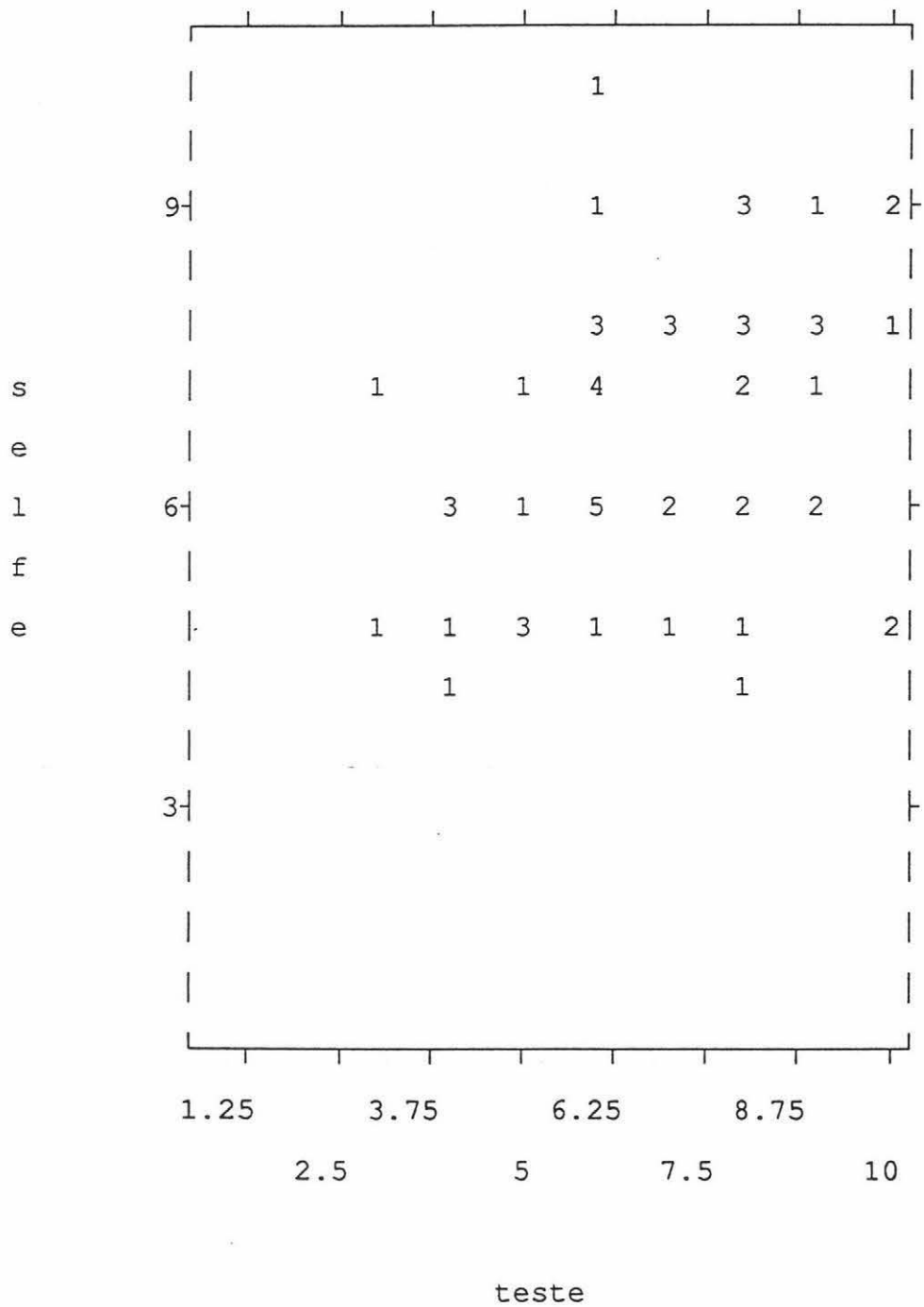
APPENDIX D.3

Scatterplot of self assessment scores and personality test scores for factor C on the 16PF.



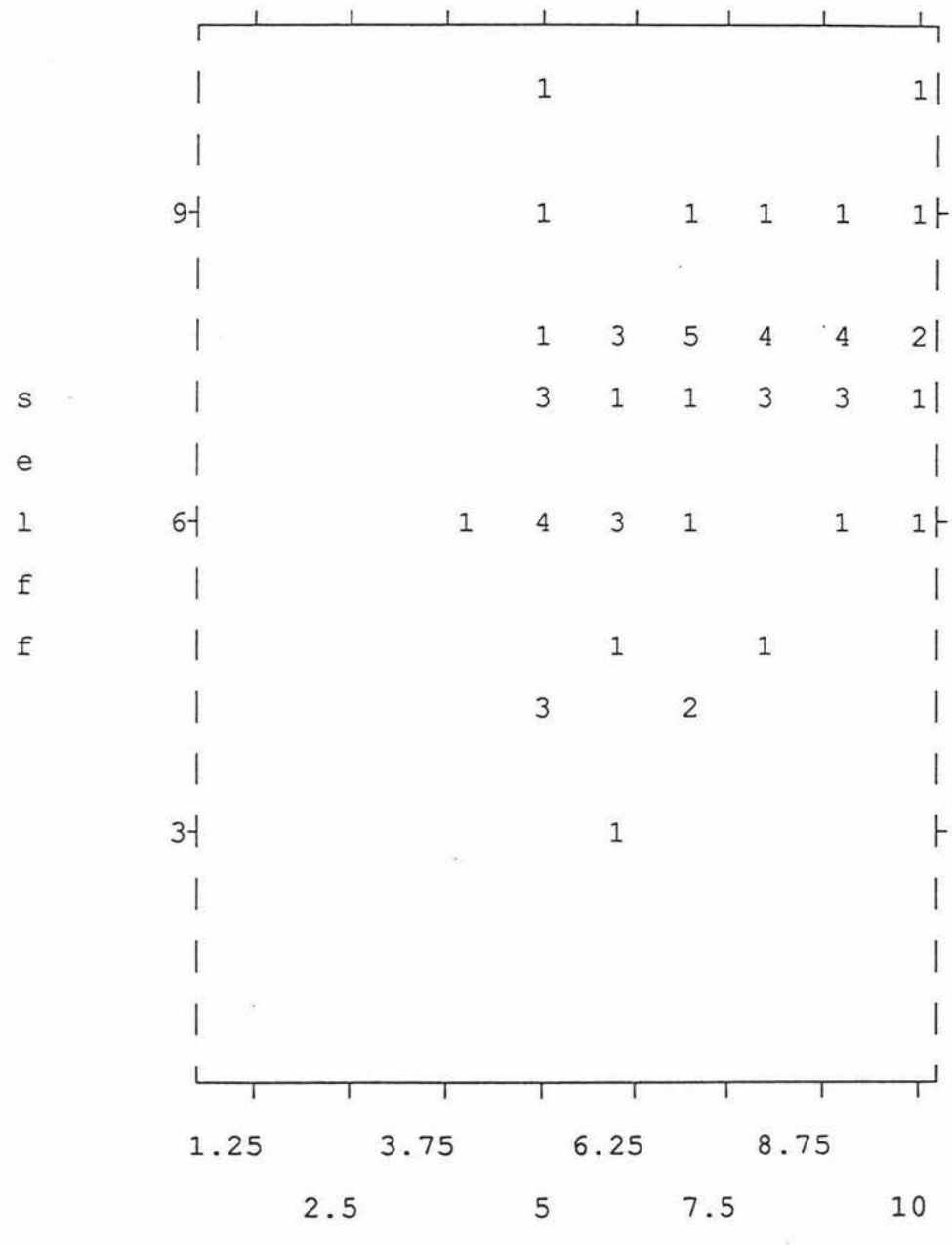
APPENDIX D.4

Scatterplot of self assessment scores and personality test scores for factor E on the 16PF.



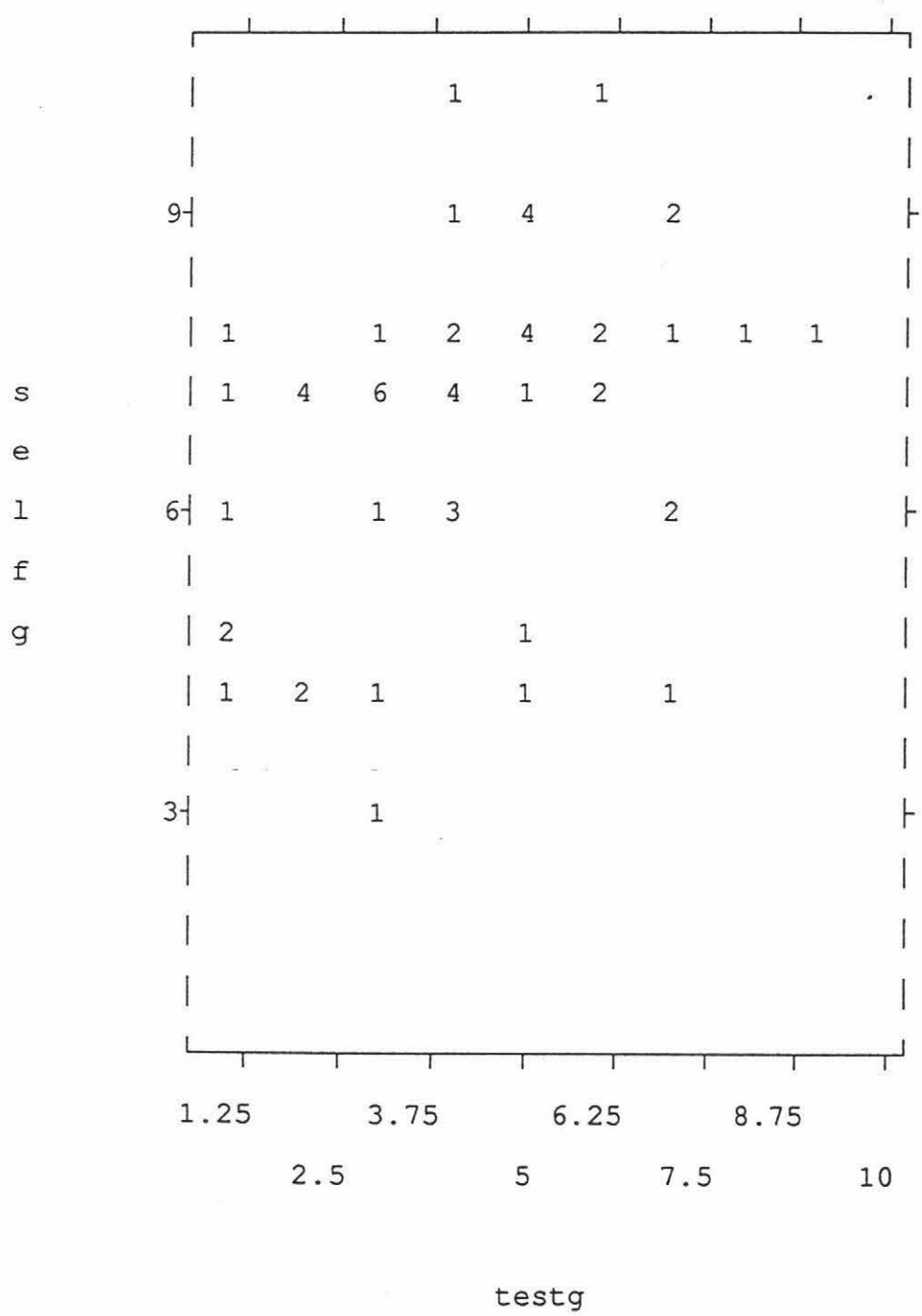
APPENDIX D.5

Scatterplot of self assessment scores and personality test scores for factor F on the 16PF.



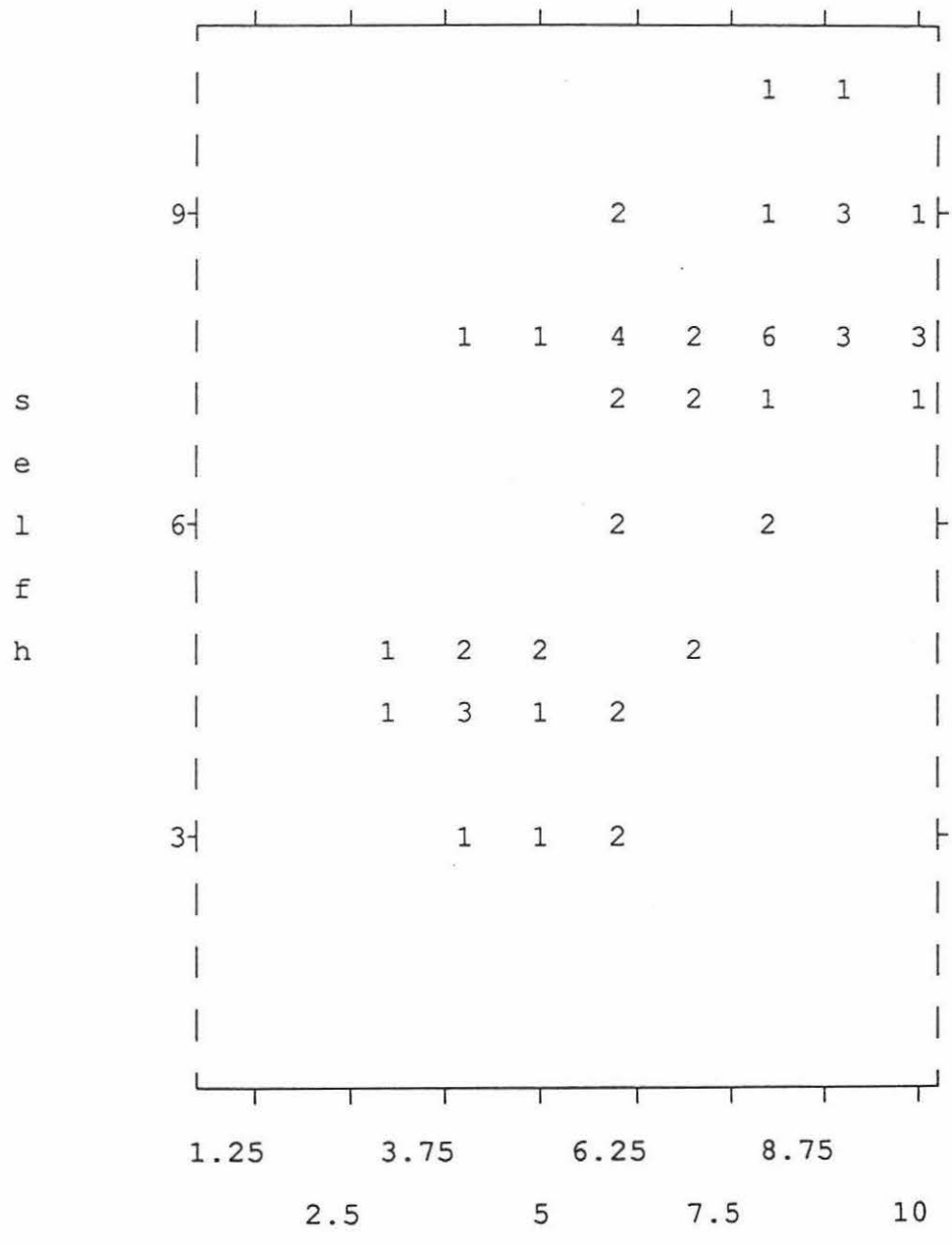
APPENDIX D.6

Scatterplot of self assessment scores and personality test scores for factor G on the 16PF.



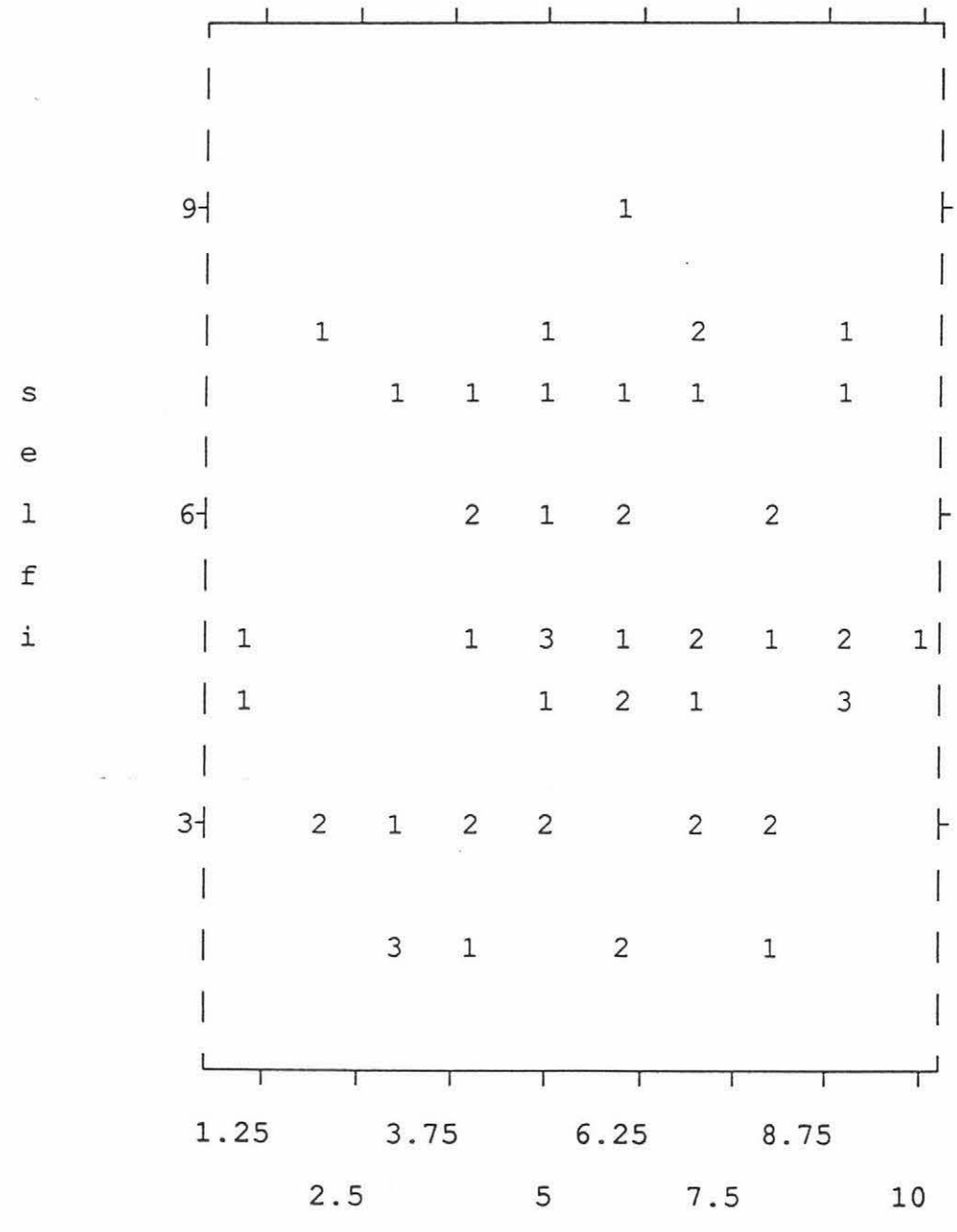
APPENDIX D.7

Scatterplot of self assessment scores and personality test scores for factor H on the 16PF.



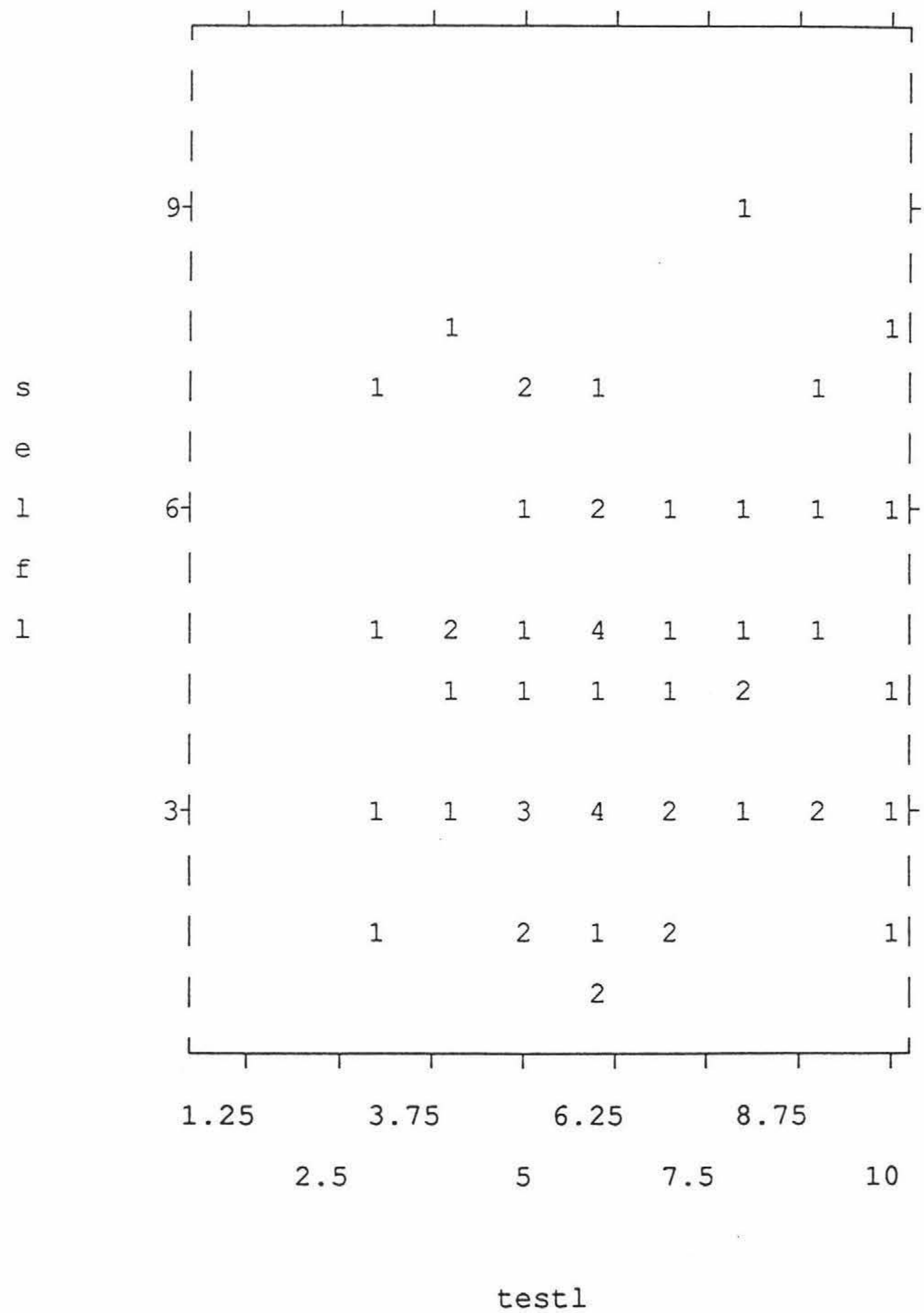
APPENDIX D.8

Scatterplot of self assessment scores and personality test scores for factor I on the 16PF.



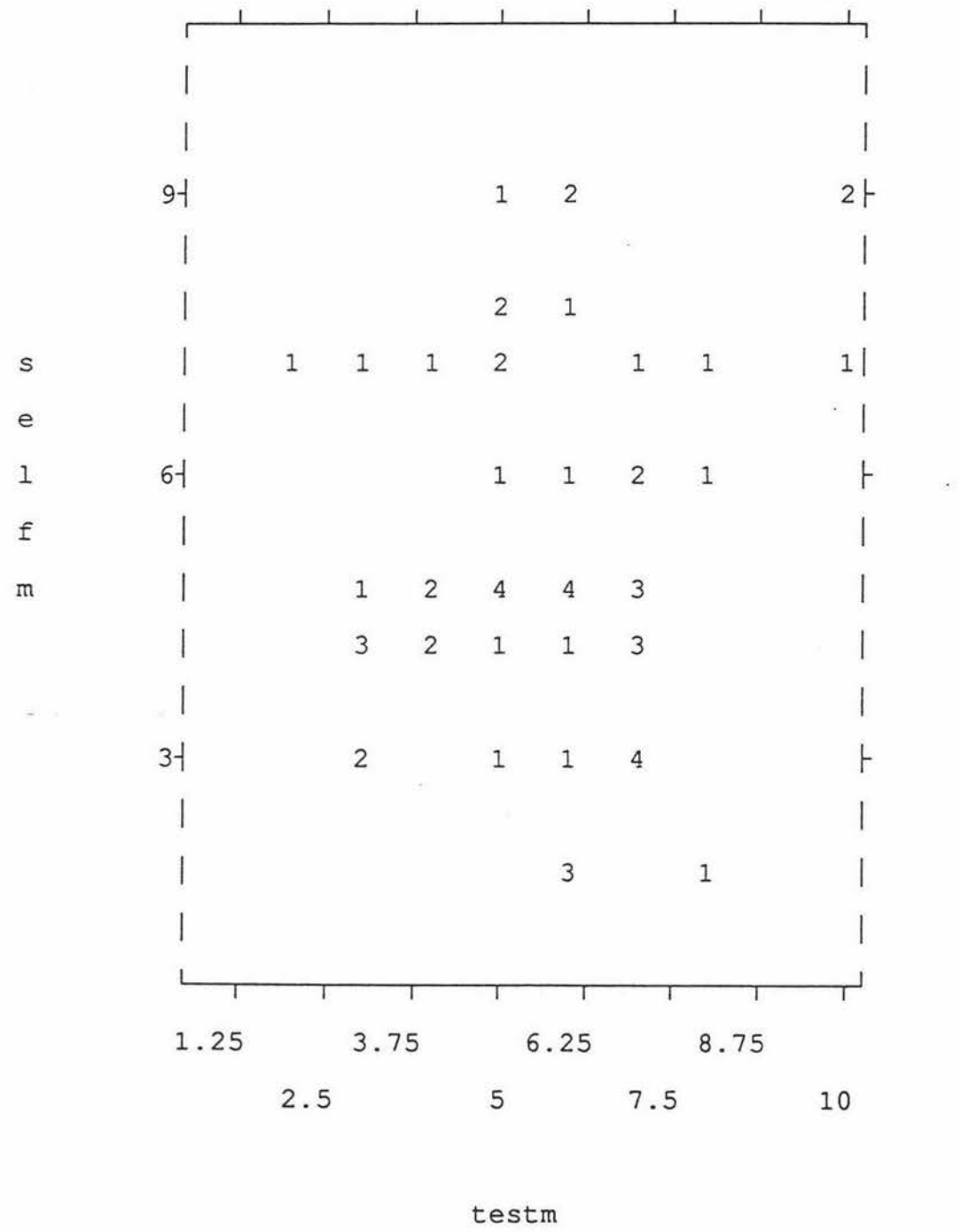
APPENDIX D.9

Scatterplot of self assessment scores and personality test scores for factor L on the 16PF.



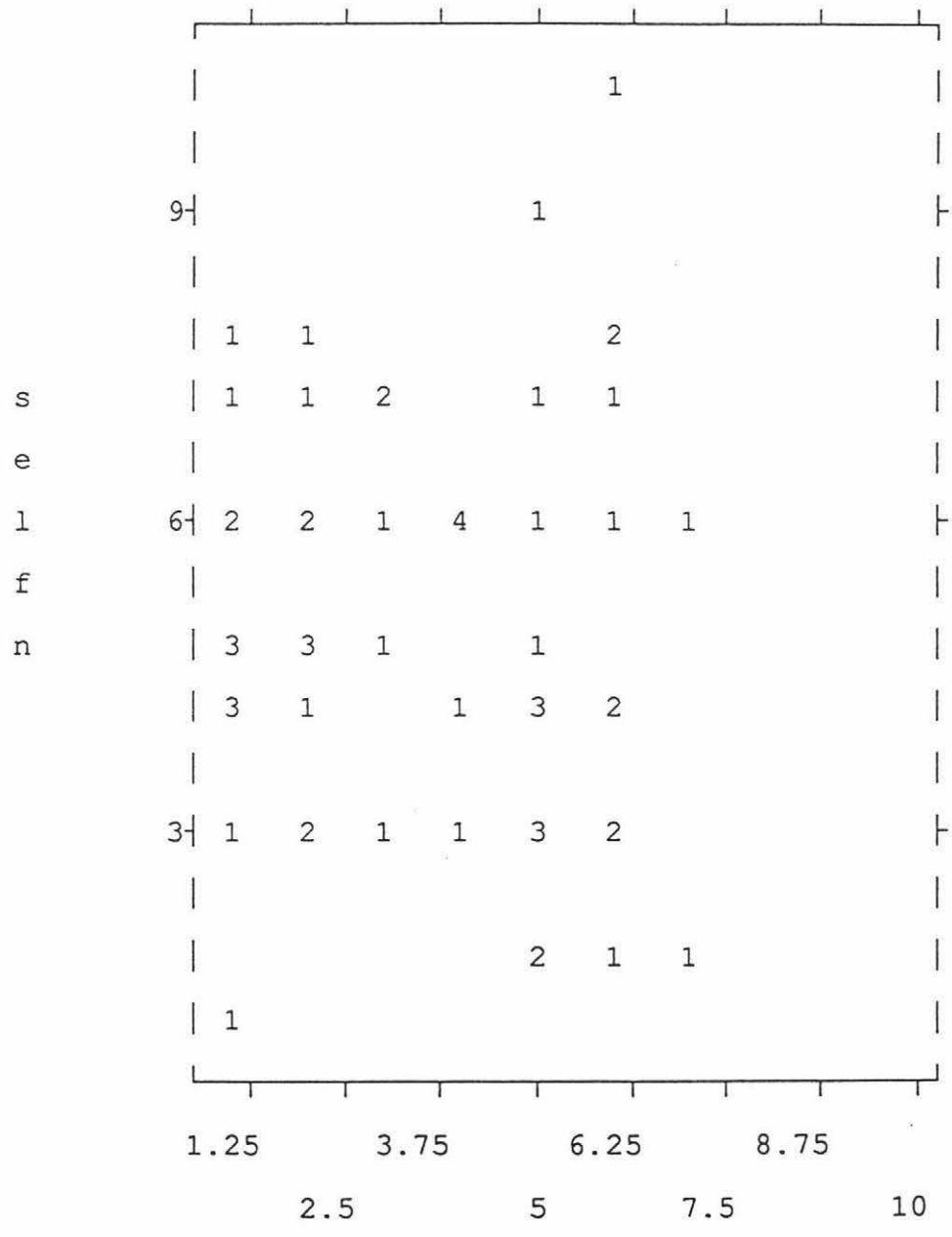
APPENDIX D.10

Scatterplot of self assessment scores and personality test scores for factor M on the 16PF.



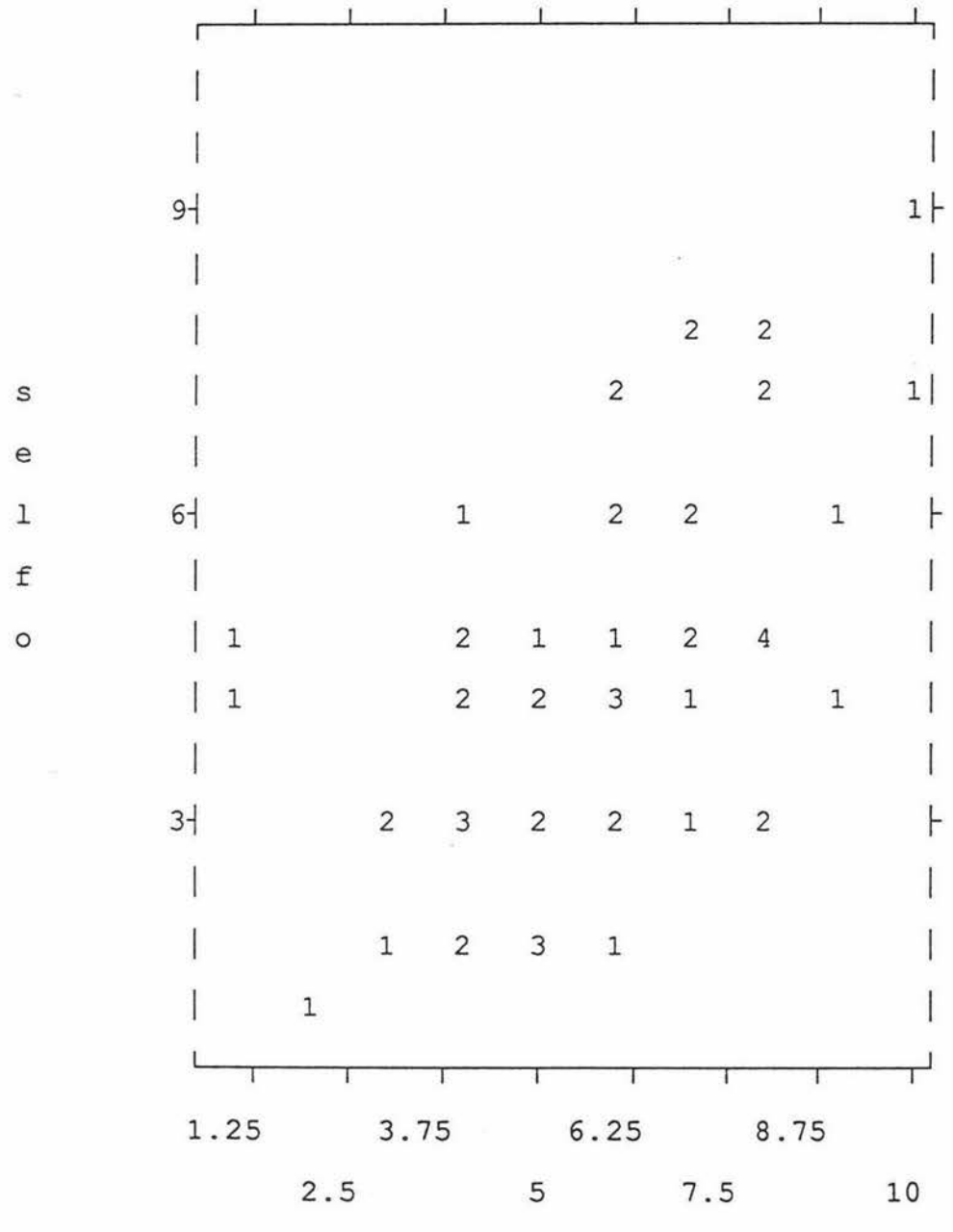
APPENDIX D.11

Scatterplot of self assessment scores and personality test scores for factor N on the 16PF.



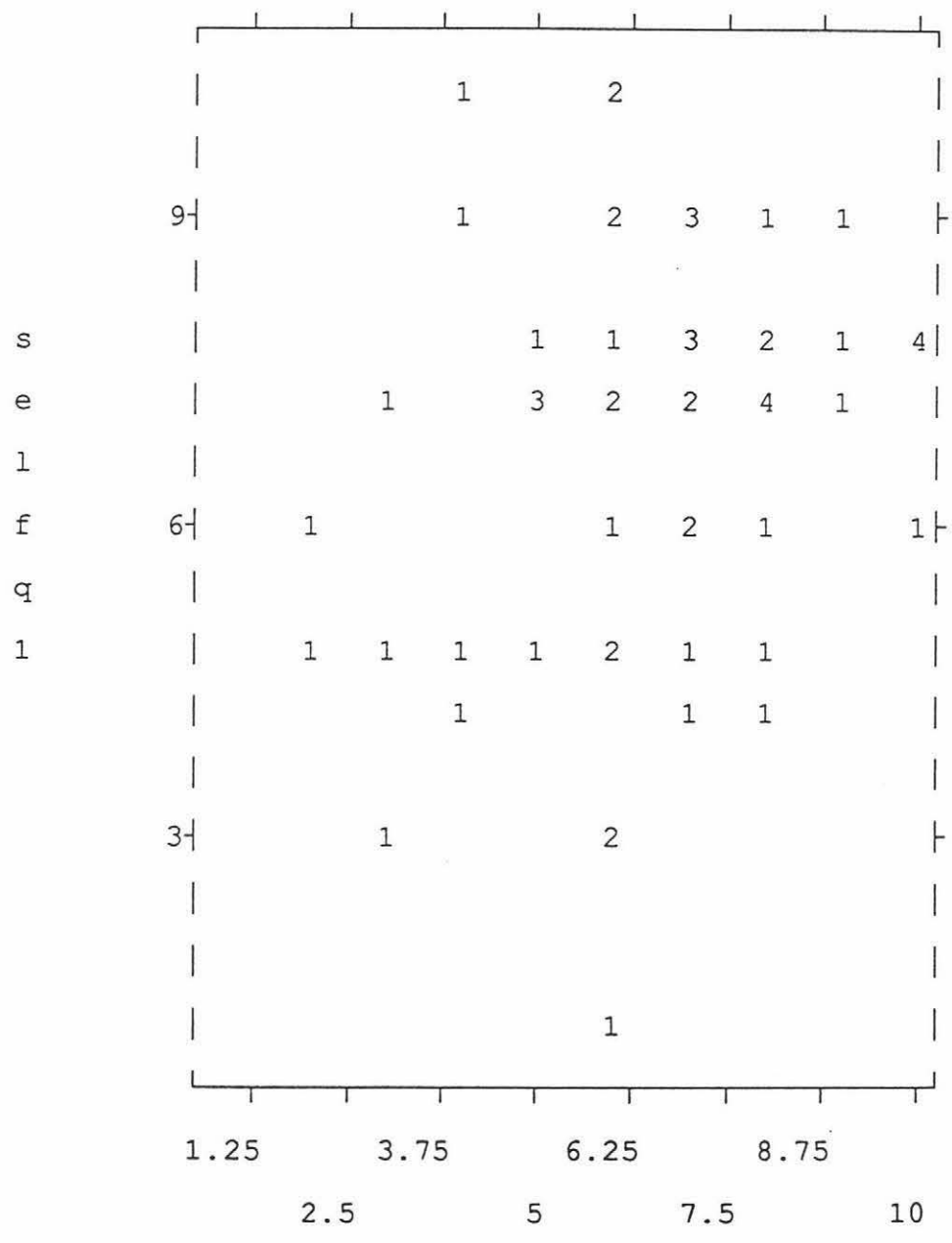
APPENDIX D.12

Scatterplot of self assessment scores and personality test scores for factor O on the 16PF.



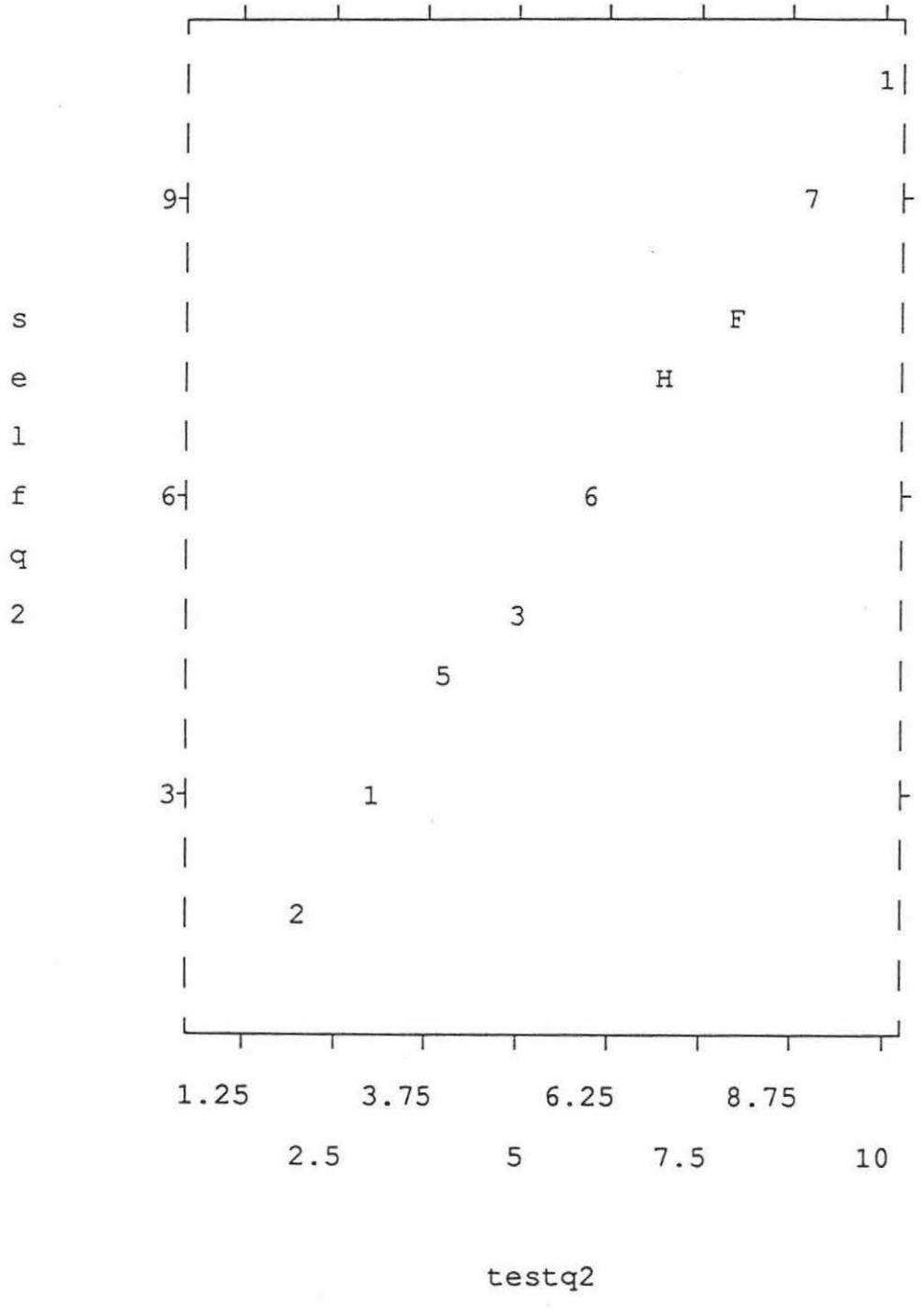
APPENDIX D.13

Scatterplot of self assessment scores and personality test scores for factor Q1 on the 16PF.



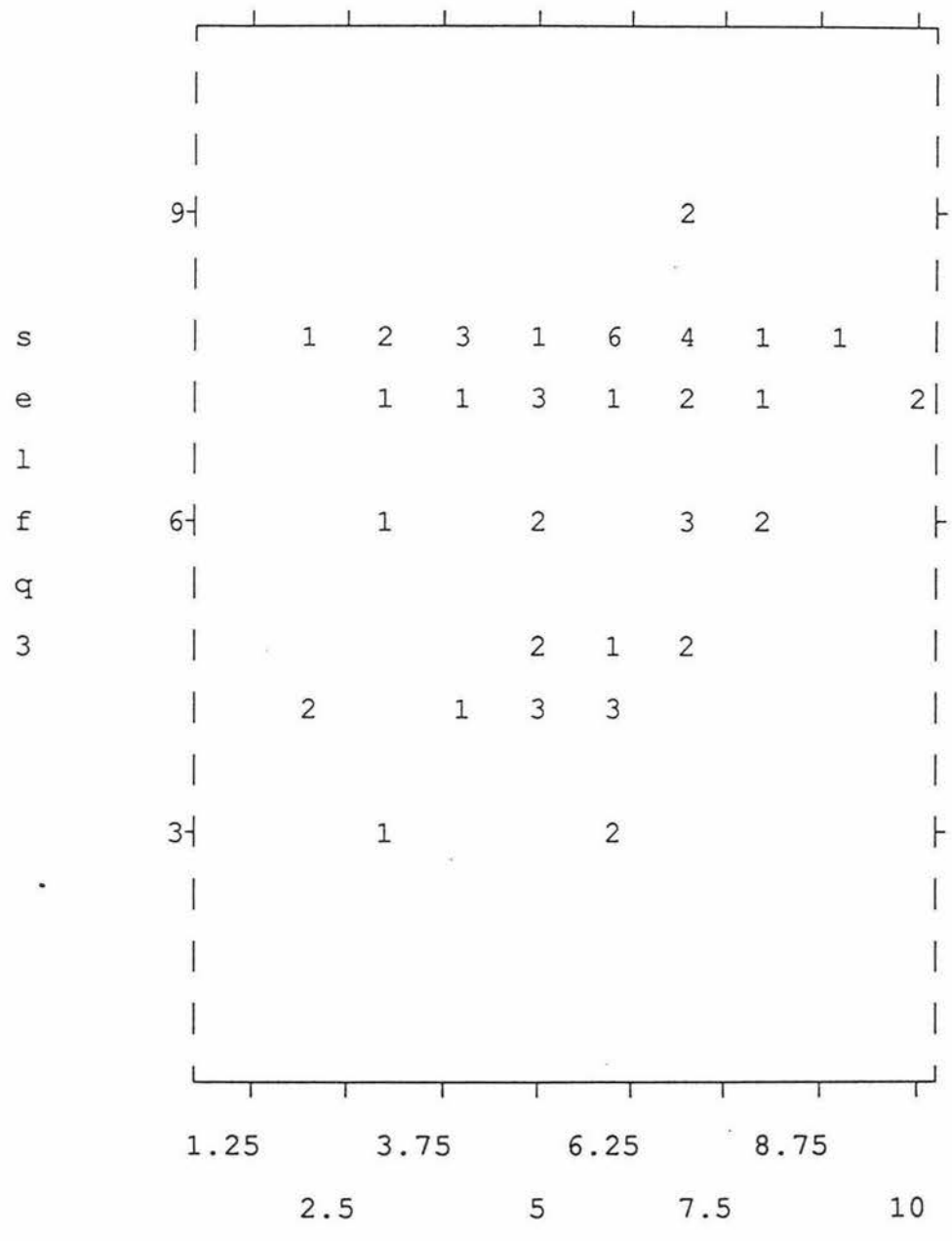
APPENDIX D.14

Scatterplot of self assessment scores and personality test scores for factor Q2 on the 16PF.



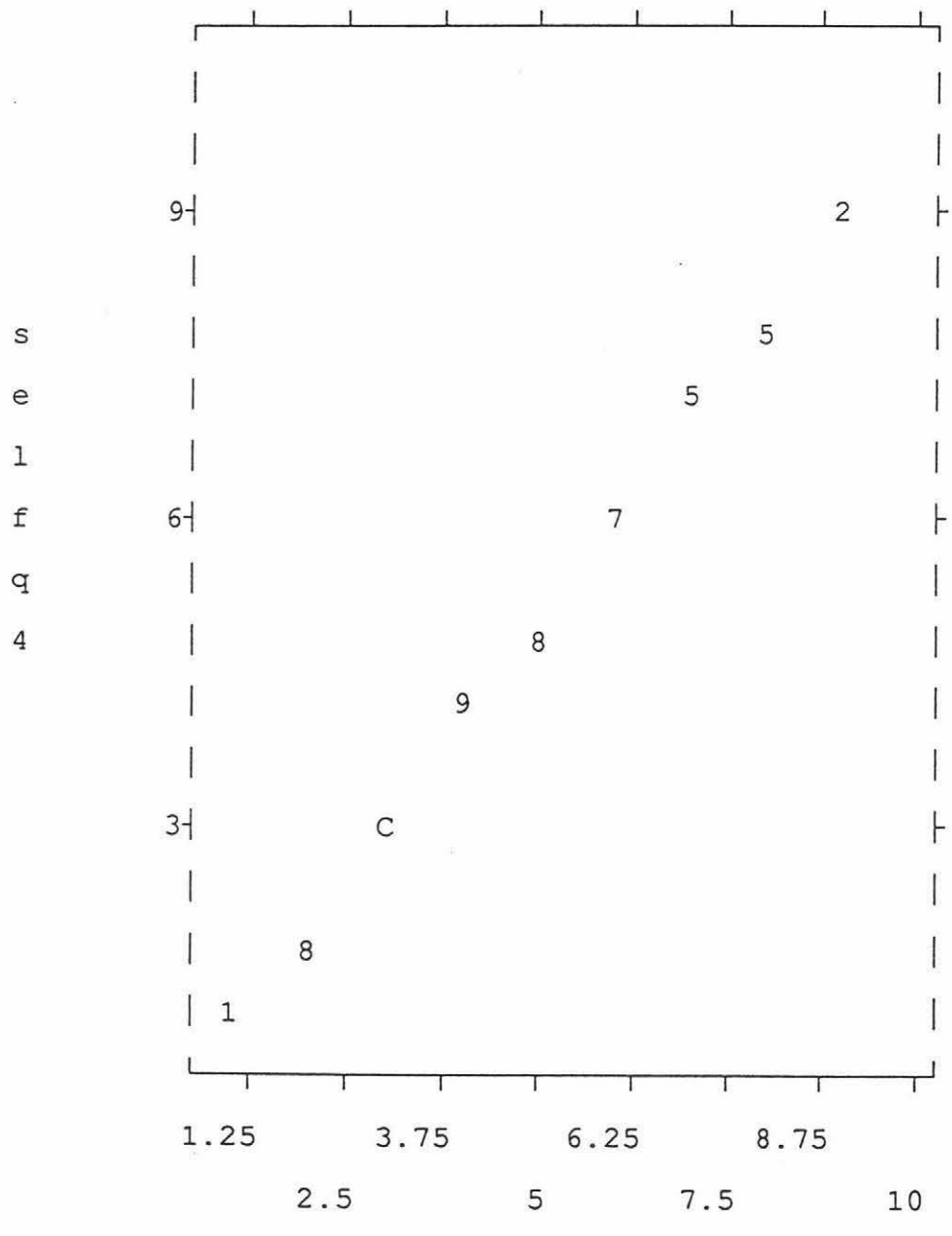
APPENDIX D.15

Scatterplot of self assessment scores and personality test scores for factor Q3 on the 16PF.



APPENDIX D.16

Scatterplot of self assessment scores and personality test scores for factor Q4 on the 16PF.



APPENDIX E

CALIFORNIA PSYCHOLOGICAL INVENTORY (CPI)
PROFILE FORM.

Profile Sheet for the CALIFORNIA PSYCHOLOGICAL INVENTORY

Female

Name _____ Age _____ Date Tested _____

Other Information _____

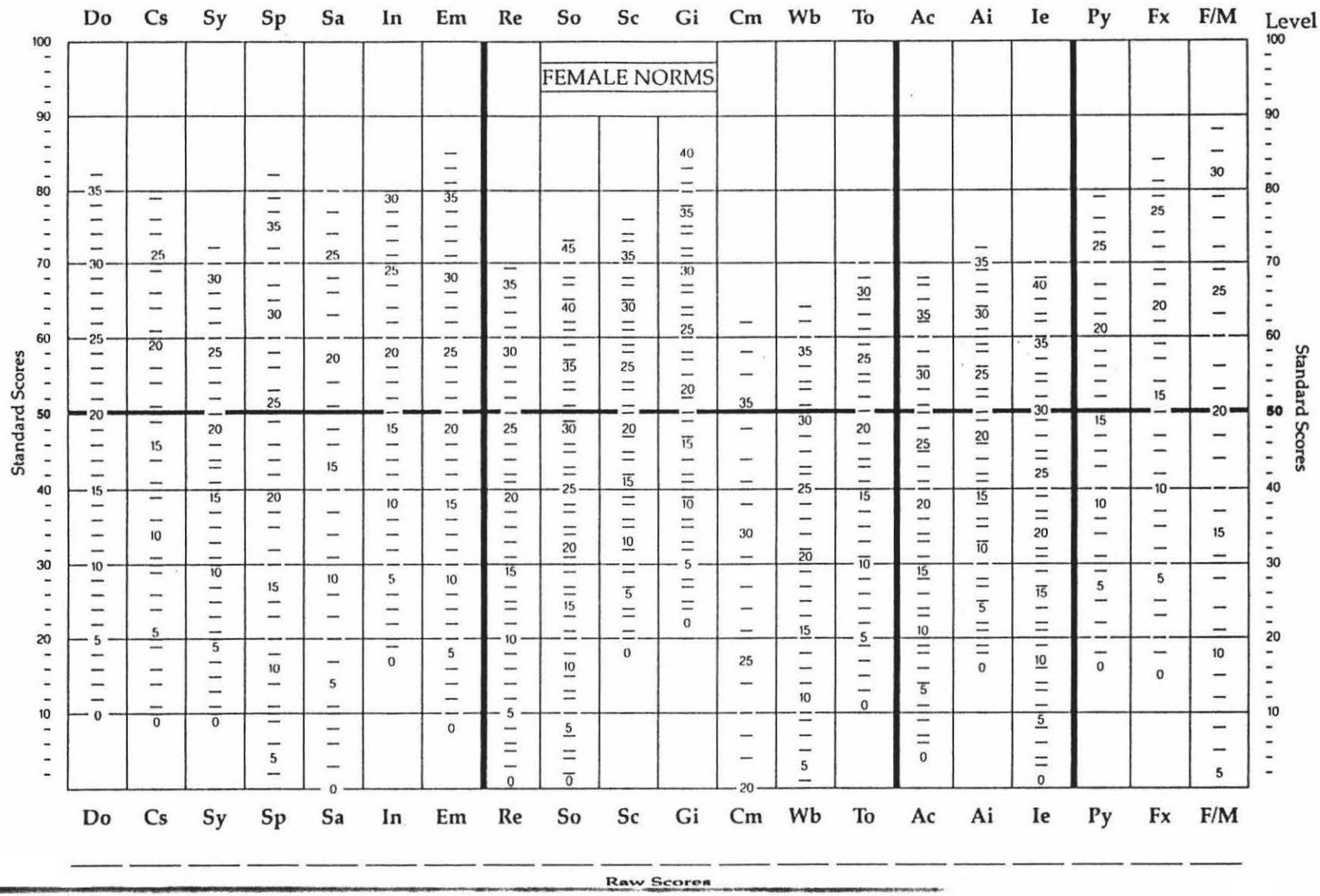
v.1 = _____ v.2 = _____

v.3 = _____

Type = _____

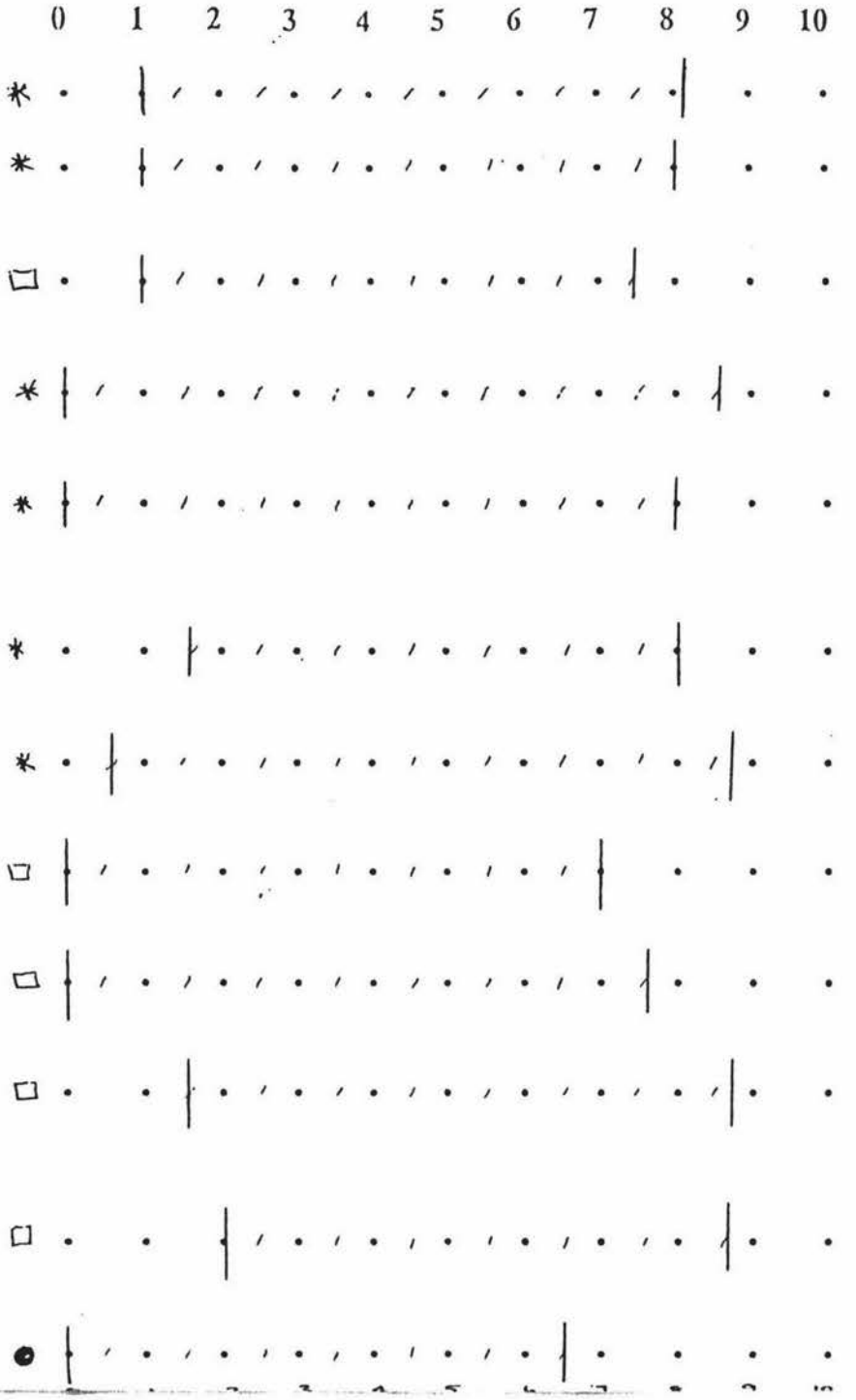
Level = _____

Notes:



APPENDIX F REVISED PROFILE FORM FOR THE CPI.

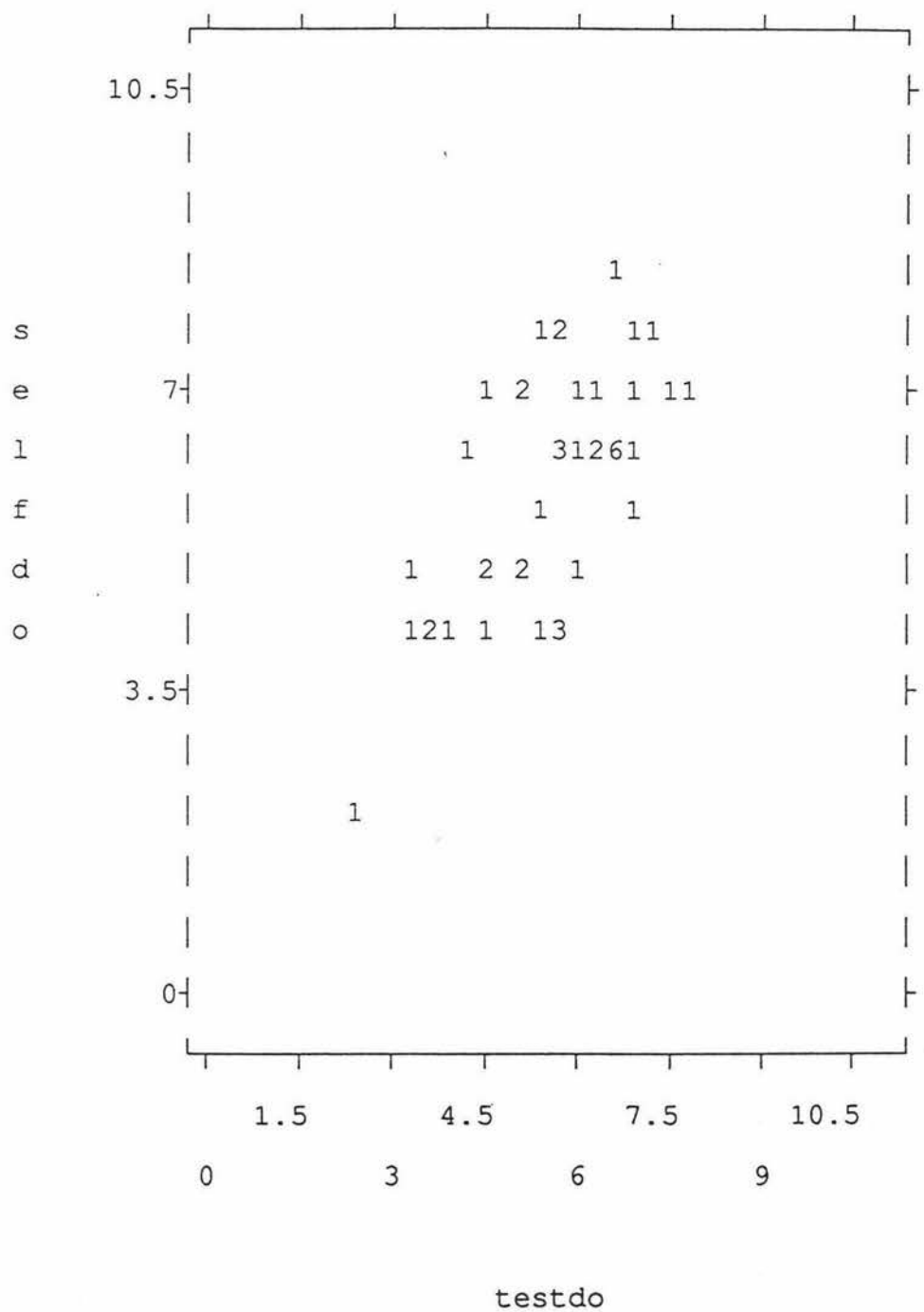
NAME: .....


0 1 2 3 4 5 6 7 8 9 10
unassuming, not forceful.
confident, assertive, dominant, task-oriented.
insure of self, dislikes direct competition.
ambitious, wants to be a success, independent.
shy, feels uneasy in social situations, prefers to keep in the background.
sociable, likes to be with people, friendly.
cautious, hesitant to assert own views or opinions, not sarcastic or sharp tongued.
self assured, spontaneous, a good talker, not easily embarrassed.
self doubting, readily assumes blame when things go wrong, often thinks others are better.
has good opinion of self, sees self as talented, and as personally attractive.
lacks self confidence, seeks support from others.
self-sufficient, resourceful, detached.
ill at ease in many situations, unempathic.
comfortable with self and well accepted by others, understands the feelings of others.
not overly concerned about duties and obligations, may be careless or lazy.
responsible, reasonable, takes duties seriously.
resists rules and regulations, finds it hard to conform, not conventional.
comfortably accepts ordinary rules and regulations, finds it easy to conform.
has strong feelings and emotions, and makes little attempt to hide them, speaks out when angry or annoyed.
tries to control emotions and temper, takes pride in being self disciplined.
insists on being him/her self, even if this causes friction or problems.
wants to make a good impression, tries to do what will please others.
sees self as different from others, does not have the same ideas, preferences etc.
fits in easily, sees self as quite an average person.

	0	1	2	3	4	5	6	7	8	9	10	
oncerned about health and personal roblems, worried about the future.	◆	├	'	•	'	•	'	•	'	•	'	feels in good physical and emotional health, optimistic about the future.
ot tolerant of others, sceptical about hat they say.	□	•	├	'	•	'	•	'	•	'	├	is tolerant of others' beliefs and values, even when different from or counter to own beliefs.
as difficulty in doing best work in ituations with strict rules and xpectations.	□	•	├	•	'	•	'	•	'	•	├	has strong drive to do well, likes to work in settings where tasks and expectations are clearly defined.
as difficulty doing best work in ituations that are vague, poorly defined, nd lacking in clear-cut methods and tandards.	□	•	•	├	•	'	•	'	•	┤	•	has strong drive to do well, likes to work in settings that encourage freedom and individual initiative.
as hard time getting started on things, nd seeing them through to completion.	□	├	'	•	'	•	'	•	'	├	•	efficient in use of intellectual abilities, can keep on at task where others might bored or discouraged.
nore interested in the practical and oncrete than abstract, looks more at what people do than what they feel or think.	✱	•	•	├	•	'	•	'	•	├	•	more interested in why people do what they do than in what they do, good judge of how people feel and what they think about things.
not changeable, likes a steady pace and well organised life, may be stubborn and even rigid.	✱	•	•	├	•	'	•	'	•	'	•	flexible, likes change and variety, easily bored by routine life and everyday experience, may be impatient, and even erratic.
decisive, action-orientated, takes the initiative, not easily subdued, rather unsentimental.	✱	├	'	•	'	•	'	•	'	•	├	sympathetic, helpful, sensitive to criticism, tends to interpret events from a personal point of view, often feels vulnerable.
	0	1	2	3	4	5	6	7	8	9	10	

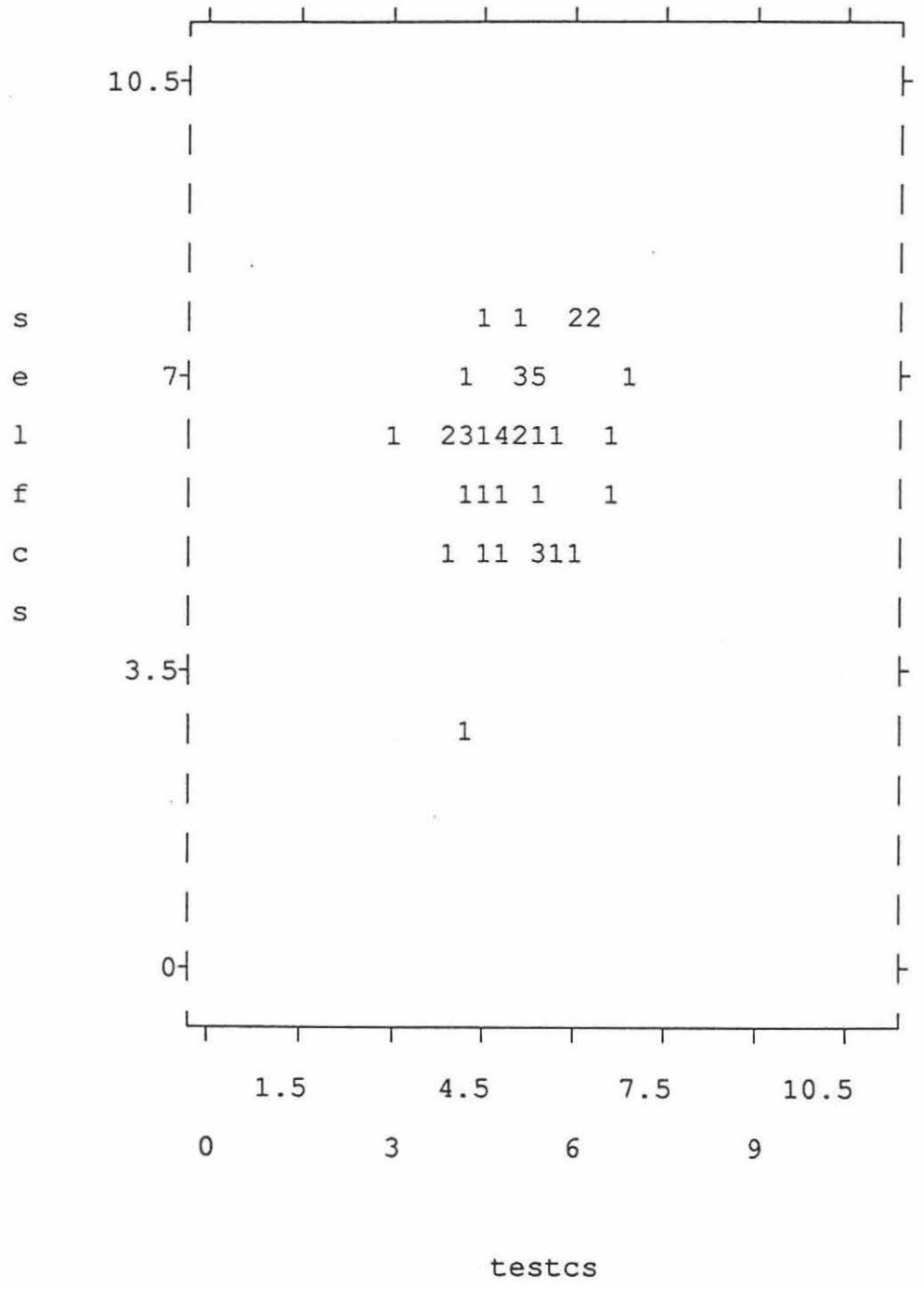
APPENDIX G.1

Scatterplot of self assessment scores and personality test scores for scale Do on the CPI.



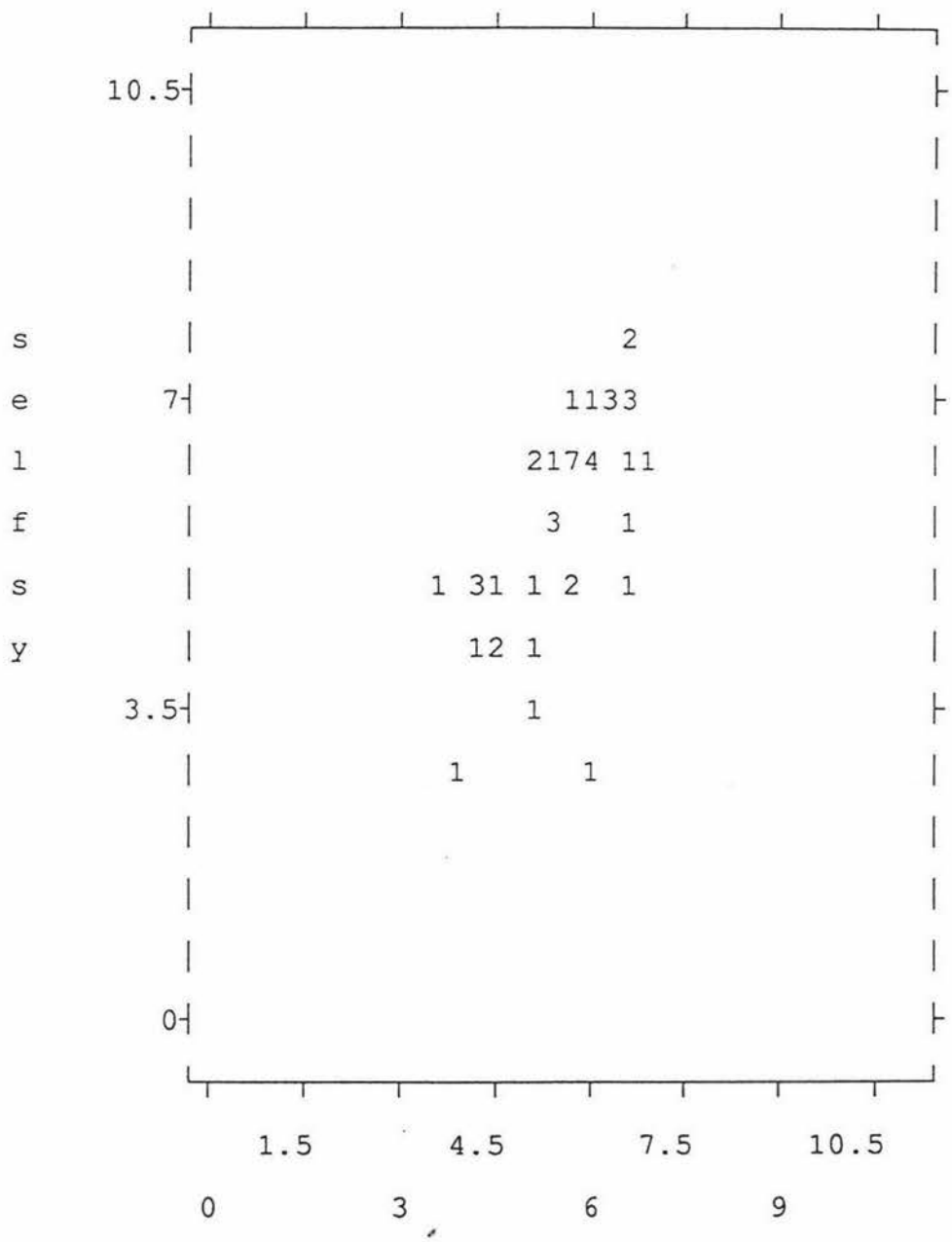
APPENDIX G.2

Scatterplot of self assessment scores and personality test scores for scale Cs on the CPI.



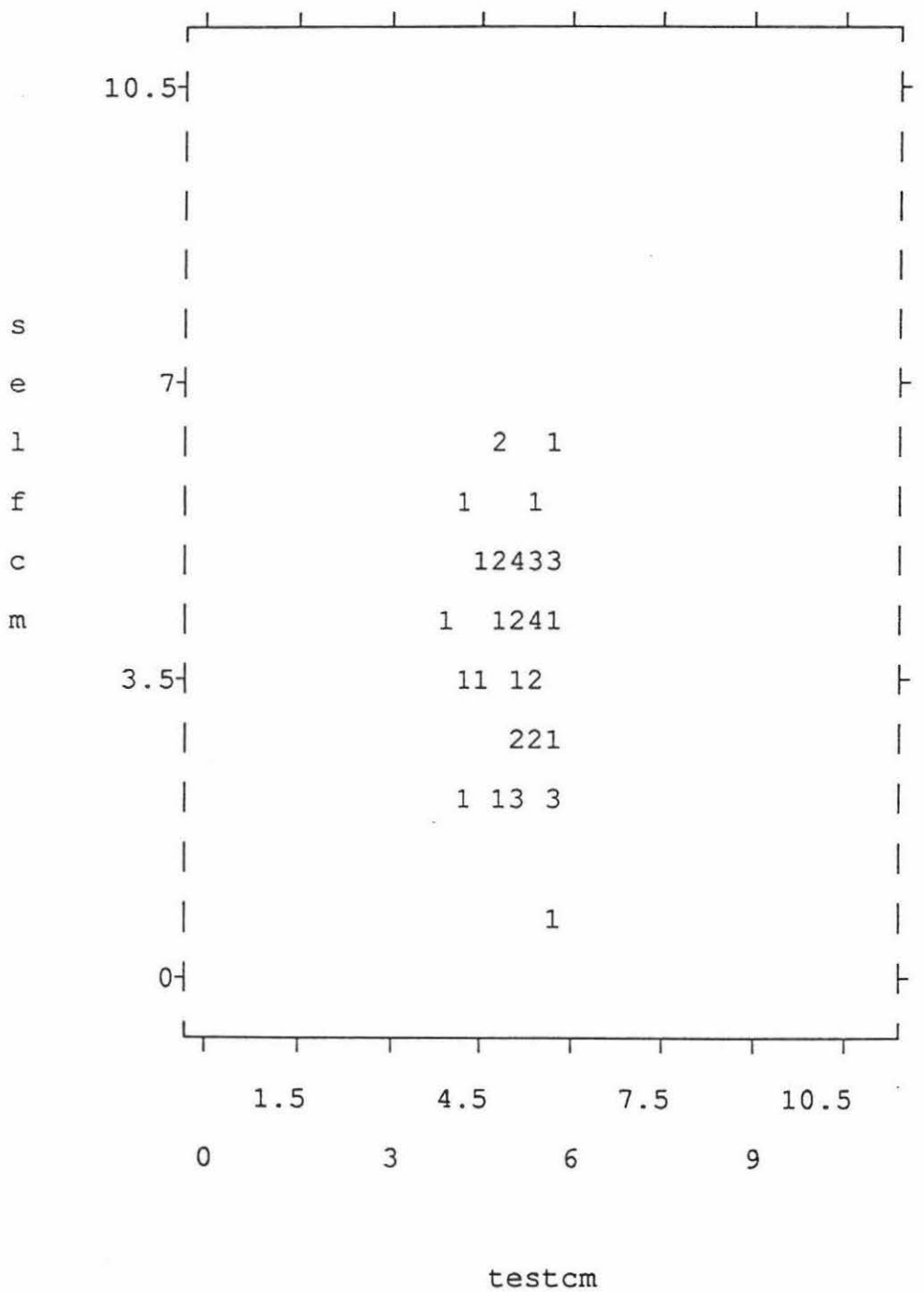
APPENDIX G.3

Scatterplot of self assessment scores and personality test scores for scale Sy on the CPI.



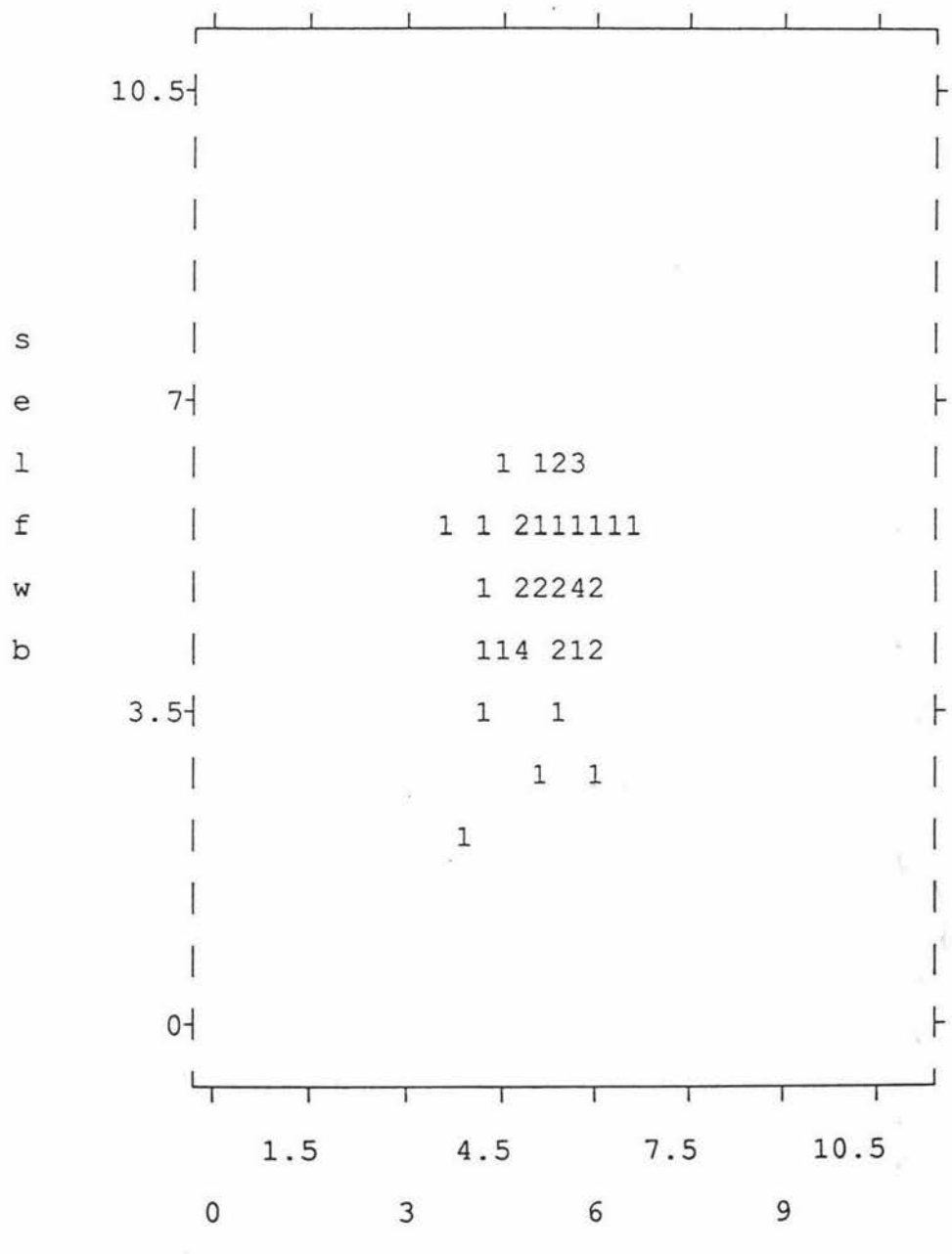
APPENDIX G.12

Scatterplot of self assessment scores and personality test scores for scale Cm on the CPI.



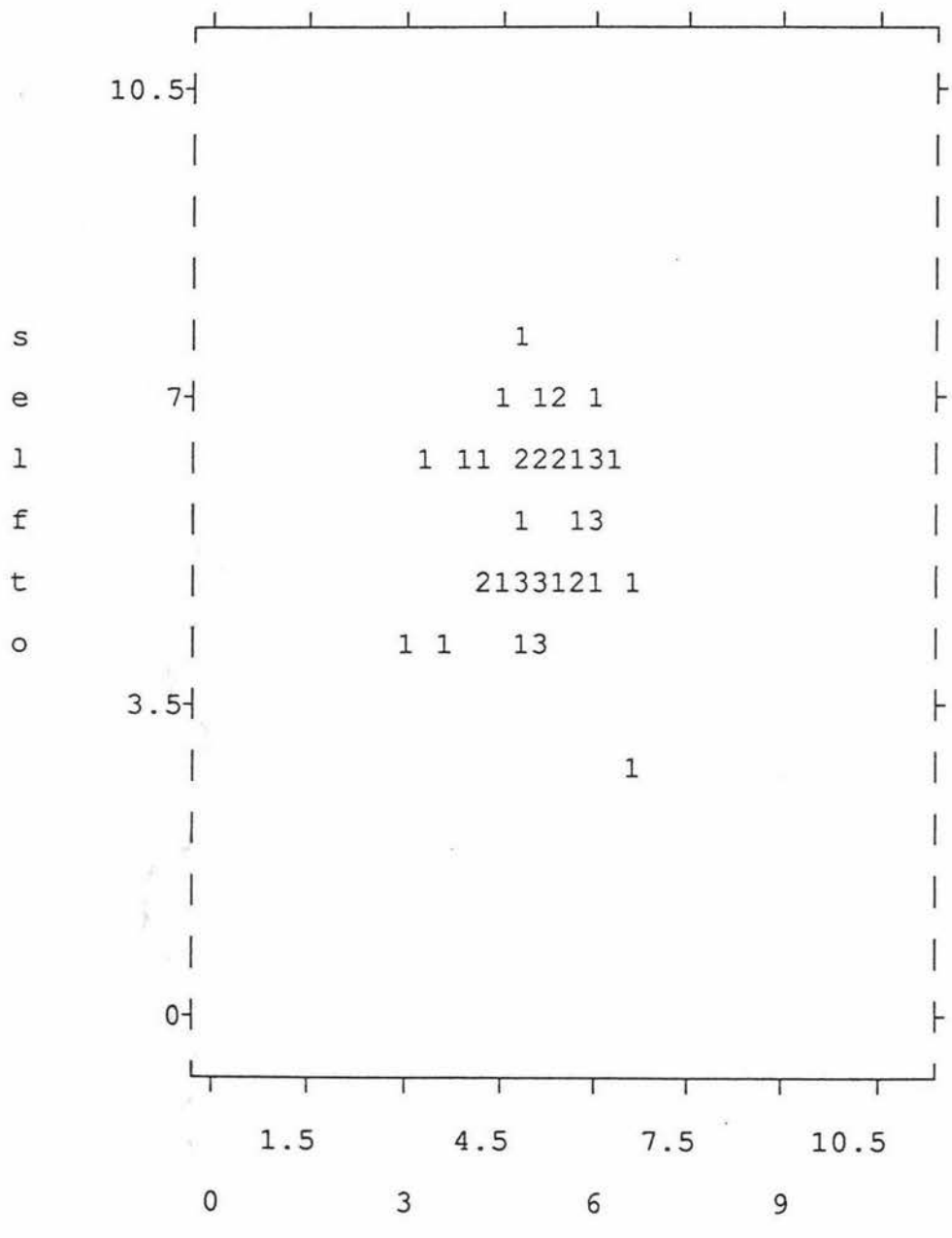
APPENDIX G.13

Scatterplot of self assessment scores and personality test scores for scale Wb on the CPI.



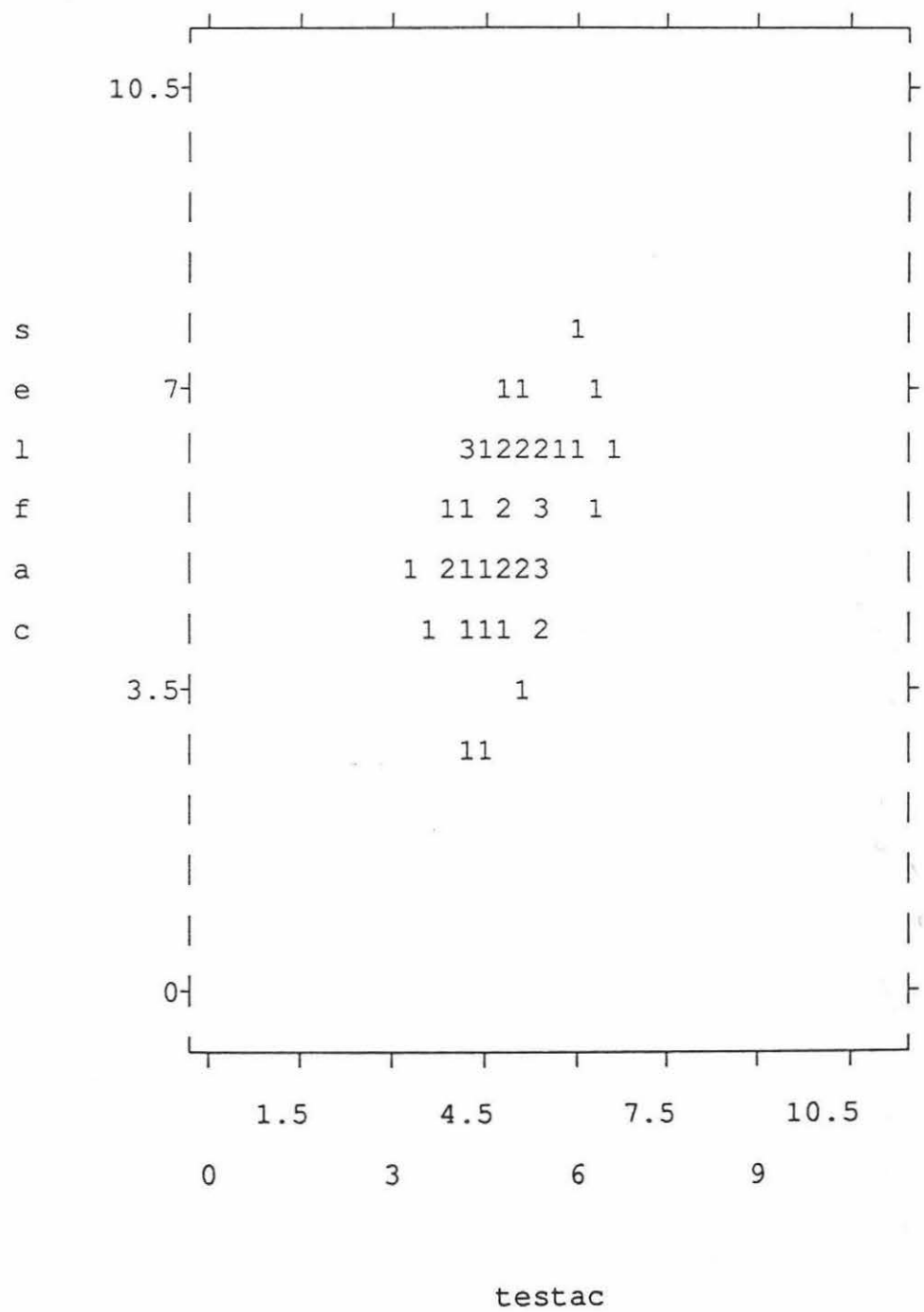
APPENDIX G.14

Scatterplot of self assessment scores and personality test scores for scale To on the CPI.



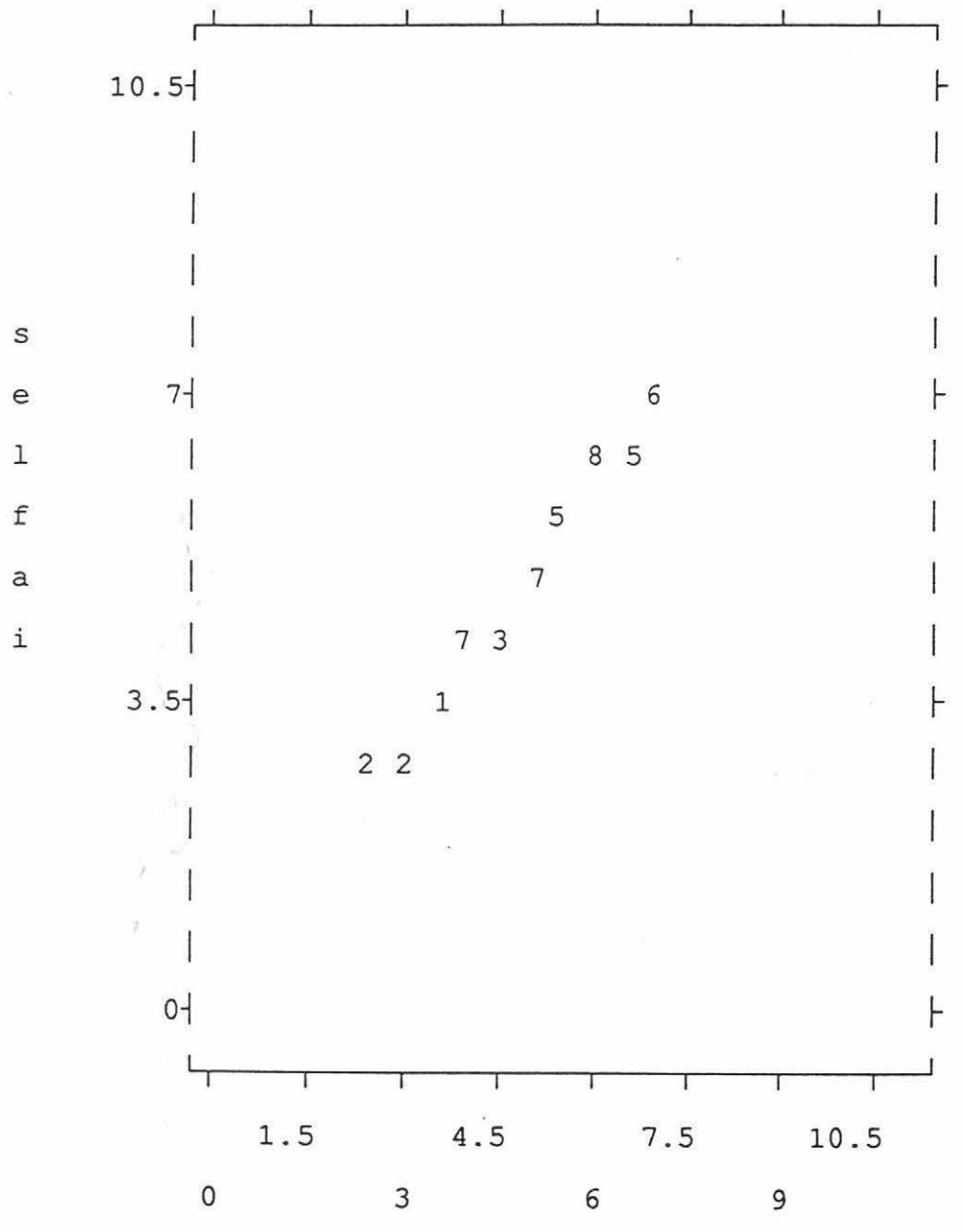
APPENDIX G.15

Scatterplot of self assessment scores and personality test scores for scale Ac on the CPI.



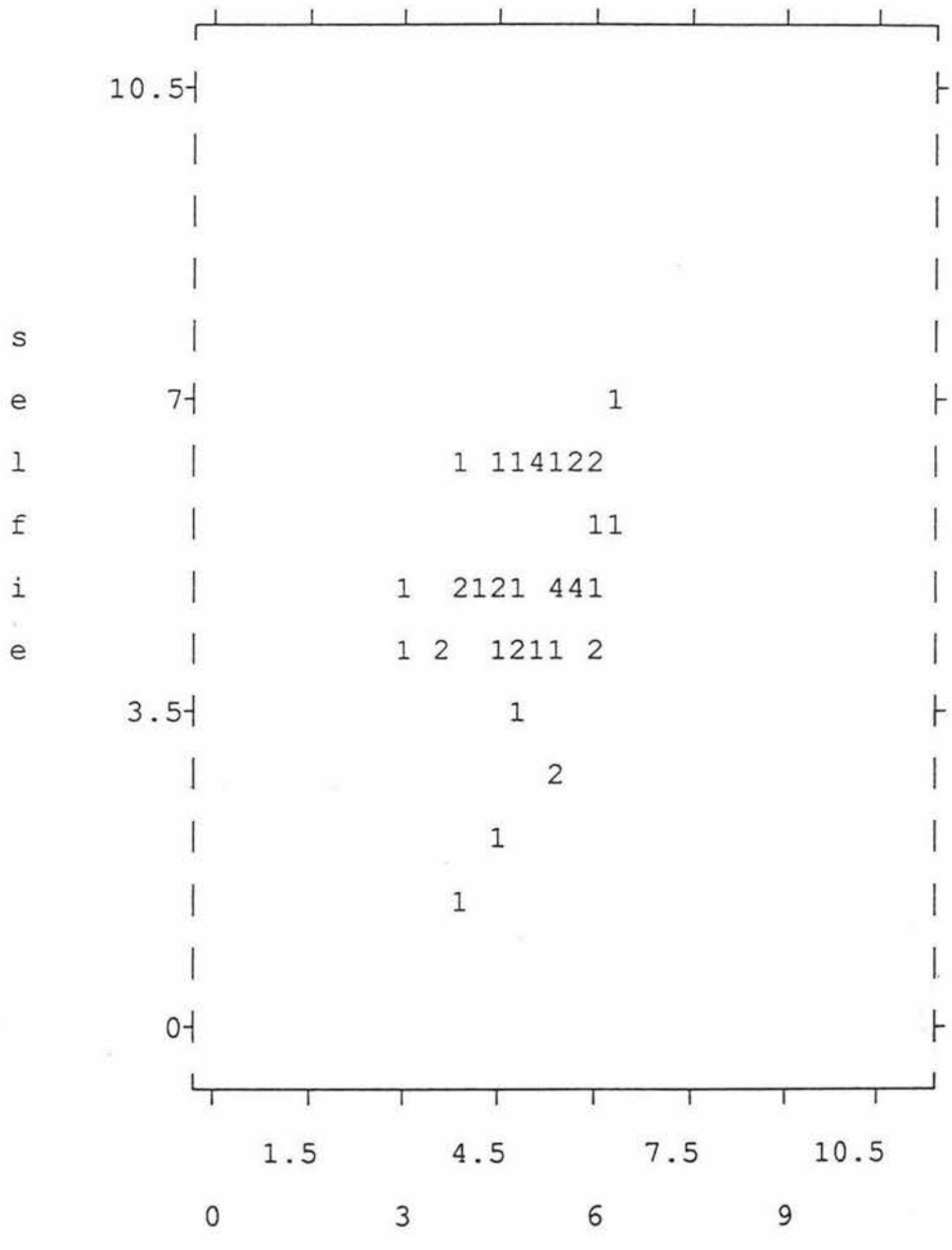
APPENDIX G.16

Scatterplot of self assessment scores and personality test scores for scale Ai on the CPI.



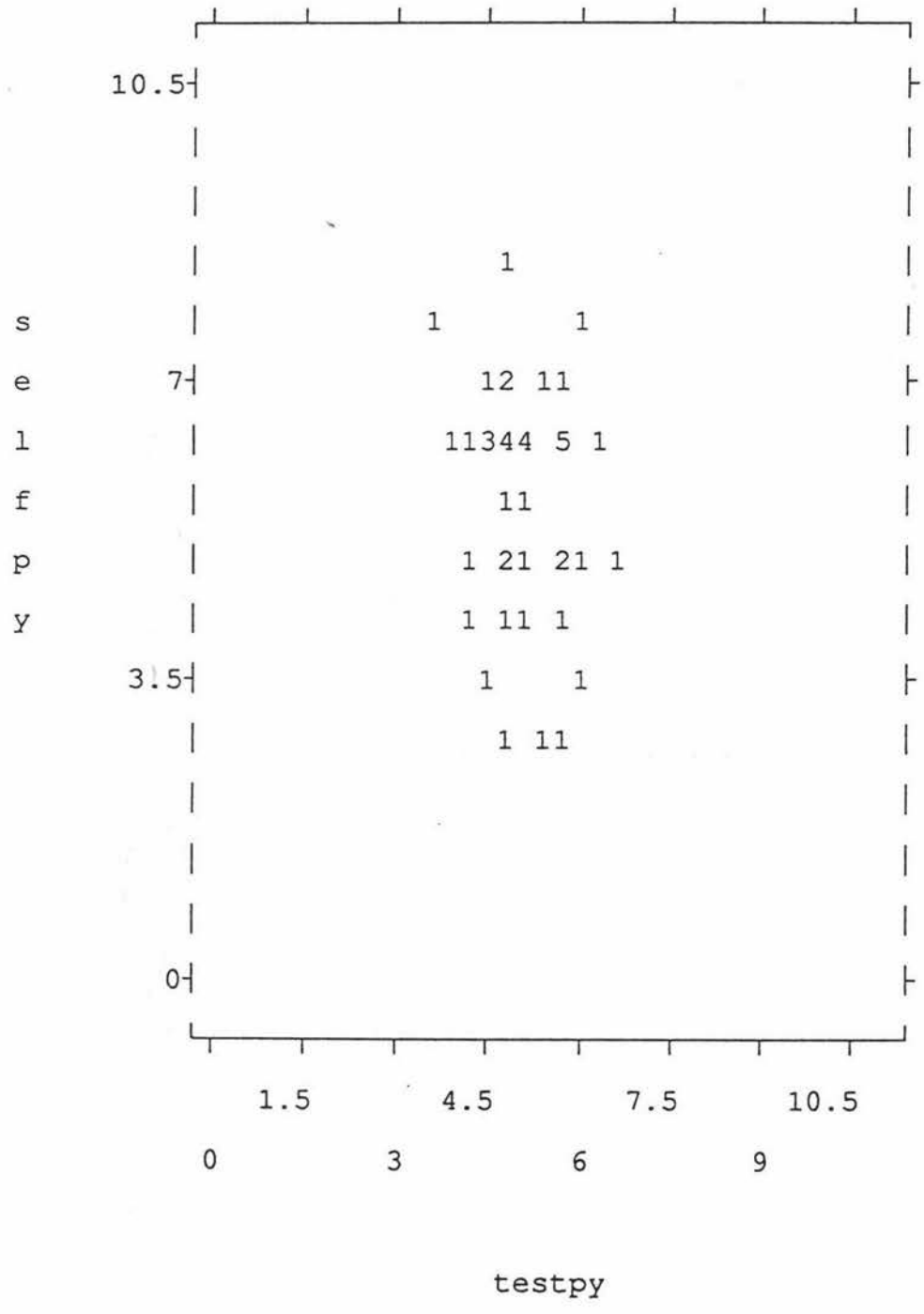
APPENDIX G.17

Scatterplot of self assessment scores and personality test scores for scale Ie on the CPI.



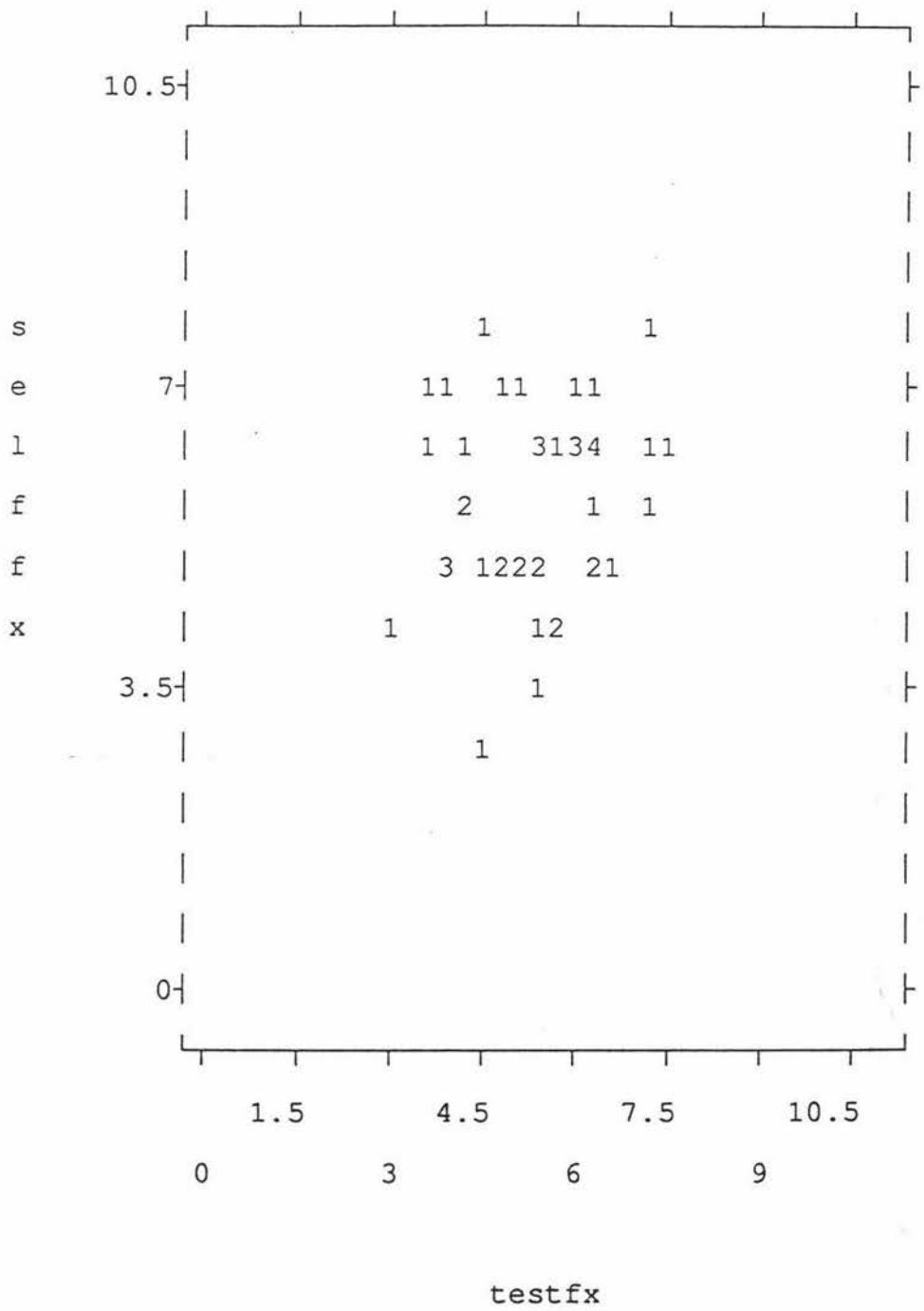
APPENDIX G.18

Scatterplot of self assessment scores and personality test scores for scale Py on the CPI.



APPENDIX G.19

Scatterplot of self assessment scores and personality test scores for scale Fx on the CPI.



APPENDIX G.20

Scatterplot of self assessment scores and personality test scores for scale Fm on the CPI.