CHAPTER 6

SOME PSYCHOMETRIC CONSIDERATIONS OF THE STUDY

6.1 INTRODUCTION

In this chapter the psychometric approach used in the study is extensively discussed. The Work Value Survey - module was constructed by the researcher. It is based on the Value Survey Module developed by Hofstede (1980a), employing well-matched samples in 40 countries around the world, and the Activism and Powerful Others - scale developed by Levenson (1974). This scale is actually a refinement of Rotter's Internal - External Locus of Control scale. The concepts of validity and reliability are discussed in detail. Reliability estimates were determined for each of the scales and are reported in this chapter.

6.2 VALUE SURVEY MODULE

Hofstede (1980a) developed this questionnaire in a large United States-based multinational corporation with subsidiaries in countries throughout the world. These subsidiaries, employing workers ranging from blue-collar labour through to those at managerial level, were organized along similar lines. Nationals of the country concerned were employed almost exclusively in the research project. Eventually 116 000 employees in 40 countries participated in the development and standardization of the questionnaire. Samples from these countries were well-matched with regard to demographical aspects. They had the same employer, similar job levels, income distributions

and came from heterogeneous age brackets. However, they differed on the concept of nationality (Hofstede, 1990, p 103). Hofstede administered a questionnaire on work values containing 120 questions, to these samples. The questionnaire was in English and the questions were derived from (Hofstede, 1980a, p 68):

1. The Allport-Vernon-Lindzey Study of Values.
2. L.V. Gordon's Survey of Interpersonal Values.
3. L.V. Gordon's Survey of Personal Values.
4. L.V. Gordon's Personal Profile.
5. G.W. England's Personal Value Questionnaire.
6. Fiedler's Least Preferred Co-Worker.
7. W.C. Schutz's Firo-B.

According to their content, these 120 questions could be divided into:

1. Questions regarding job satisfaction.
2. Questions regarding work behaviour.
3. Questions in respect of personal goals and beliefs.
4. Demographical questions.

(Hofstede 1980a, p 66).

Responses were evaluated by means of a five-point Likert scale. In order to test for convergence and correlation of the scores, Hofstede simultaneously administered his questionnaire plus value scales developed by Wollack, Super and Rosseel, as well as personality tests. Data was extensively analysed by means of analysis of variance and factor analysis. The factor analysis yielded, besides six factors relating to job content, reward, interpersonal relations, company, security and comfort, the four dimensions of power distance, individualism, uncertainty avoidance

and masculinity. These four dimensions emphasize the differences in national cultures (Hofstede, 1990, p 104). Hofstede refined the instrument which resulted in the Value Survey Module incorporated in the Work Value Survey Module (See appendices). Hermann (1989) administered Hofstede's Value Survey Module to a multi-ethnic group of respondents. He determined a reliability estimate and obtained a split-half coefficient with Spearman-Brown correction for unequal length of 0,9749 and an alpha coefficient of 0,9913 for part one and 0,9172 for part two.

6.3 ACTIVISM AND POWERFUL OTHERS-SCALE

Rotter and his associates (1966) developed the concept of Internal-External Locus of Control. They employed it to study the effect of reward on behaviour. An internally orientated person believes that his/her own behaviour affects the rewards which follow on it. An externally controlled person believes that outside forces shape and reward his life (Gurin, Gurin, Lao and Beattie, 1969, p 29). Rotter's Internal-External Locus of Control-scale (I-E scale) measures the extent to which people believe they are in command of their lives, exercising control (internally controlled) or the degree to which they feel their lives are determined by fate, chance or powerful others (externally controlled). As a number of closely related concepts had become prominent in the studies of low income and minority groups (Gurin et al, 1969, p 30), Levenson (1974) undertook further research to refine the I-E scale to become a valid instrument to obtain conceptually clearer measures of locus of control. Levenson (1974) separated the unidimensional locus of control scale into the three dimensions of internality, powerful others and chance. She subjected it to a validity study by means of a factor analysis to

ascertain the validity of this separation. The factor analysis yielded proof that this tripartite separation of expectations of control, adds conceptually and empirically to the usefulness of the concept of locus of control (Levenson, 1974, p 382). The Activism and Powerful Others-scale is also incorporated in the Work Value Survey Module.

6.4 BIOGRAPHICAL QUESTIONNAIRE

A biographical questionnaire forms part of the Work Value Survey Module. The questions relate to aspects such as gender and age brackets, home language, years of formal schooling received, religion, educational level, occupational level, income bracket, country of origin, ethnic group and sector of the economy employed in. Information obtained in this way provides either independent or nuisance variables in the processing of data.

6.5 VALIDITY

Bohrnstedt and Knoke (1988, p 12) define validity as the degree to which an operation results in a measure that accurately reflects the concept it is intended to measure. Babbie (1989, p 98) views validity as a descriptive term used of a measure that accurately reflects the concept it is intended to measure. Mason et al (1989, p 420) see validity as "the degree to which a test measures what it is supposed to measure". De la Rey's view (1978, p 30), that a test is valid only if it measures the concept or characteristic it pretends to measure, ties in with these definitions. Validity is usually determined by means of correlational statistics and expressed as a validity coefficient. There is also a non-statistical approach to the determination of psychological test validity, namely content validity which is a matter of judgement and not of empirical correlation (Guion, 1965, p 125).

The validity estimate is usually determined by calculating the correlation between performance in a test and an independent, objective criterion of the behaviour being measured (Smit, 1983, p 47).But this is only one kind of validity, i.e. predictive validity which could either be concurrent or predictive prediction, as is illustrated later on in this chapter. De la Rey (1978, p 31) distinguishes between construct validity, content validity, criterion-related validity, concurrent validity, face validity and synthetic validity. Construct validity is the extent to which a test measures the construct it was designed to measure (Mason et al, 1989, p 260). Construct validity is determined by comparing a new test with existing valid tests measuring the same concept. A high significant correlation points to construct validity (Smit, 1983, pp 63-67). Construct validity evaluates the construct as well as the adequacy of the test in measuring the construct (Mason et al, 1989, p 261; Smit, 1983, p 64). Dane (1990, p 259) and Smit (1983, p 66) distinguish three approaches to the study of construct validity, viz convergent validity, discriminant validity and factorial analysis. Convergent validity points to the extent to which a measure correlates highly with existing psychological tests measuring the same concept. Discriminant validity, on the contrary, is the extent to which a measure does not correlate too obviously or not at all with tests measuring different concepts. The construct discriminates between similar and entirely different constructs (Smit, 1983, p 66). By means of factor analysis, the number of common factors, explaining the variance, are identified. These factors can predict performance in a test. By identifying the factors common to a construct, it is possible to construct a test which is a refined and clear measure of a specific theory or concept (Smit, 1983, p 66).

Content validity is of a qualitive nature and ascertains the degree of representativeness of the contents of a questionnaire of the construct being measured (De la Rey, 1978, p 31). Criterion-related validity may be separated into predictive validity and concurrent validity (Howard, 1985, p 100). Predictive validity concerns the degree to which a test predicts future behaviour or performance correctly (Smit, 1983, p 51). A predictive validity estimate is determined by means of Bravais-Pearson product moment correlation or multiple regression analysis (De la Rey, 1978, p 31). The validity coefficient is usually interpreted by way of its numerical size (magnitude), coefficient of alienation, coefficient of determination and the standard error of measurement (Smit, 1983, pp 52-53). Concurrent validity implies the degree to which test variance correlates with variance in a test (criterion) available at essentially the same time (Smit, 1983, p 61). Smit views concurrent validity as a relationship expressed in terms of a correlation coefficient between a test score and another yielded by a measure already accepted as valid of the same behavioural construct (1983, p 62). In other words, concurrent validity involves comparing a new measure to an existing valid measure with the emphasis on the present status of the measure or the respondent (Smit, 1983, p 62). Face validity or expert validity is the degree of consensus between experts that a measure represents a particular concept (Dane, 1990, p 257). Synthetic validity refers to presumed validity (De la Rey, 1978, p 31). Howard (1985, p 56) also distinguishes between external and internal validity. External validity deals with the extent to which a researcher can generalize across samples, situations, settings and times based on evidence from a particular study. Internal validity is defined as the extent to which procedures enable one to draw reasonable conclusions (Howard, 1985, p 110).

6.5.1 VALIDITY OF THE WORK VALUE SURVEY MODULE

Construct validity is of primary importance here. Hofstede (1980a) determined the construct validity of his value survey questionnaire by means of a factor analysis yielding the four value dimensions of individualism, masculinity, uncertainty avoidance and power distance. Factorial analysis done on data procured by the repeated application of Hofstede's Value Survey Module yielded the same results (Hofstede, 1980b; Hofstede and Bond, 1984; Singh, 1990). In the case in hand the data was submitted to a principal axis factoring with varimax rotation. The obtained results differed completely from previous validation studies. After rotation only two factors with eigenvalues greater than one could be obtained. The eigenvalues are presented in Table 6.1.

Table 6.1: EIGENVALUES: EXTRACTED FACTORS - VALUE SURVEY MODULE.

Factor	Eigenvalue	Percentage of variance	Cumulative Percentage
1	8,58	31,8	31,8
2	1,23	4,5	36,3
3	0,79	2,9	39,2
4	0,65	2,4	41,7

Table 6.1 shows that the factors with eigenvalues greater than one declare only 36,3% of the variance.

However, the rotated factor matrix which is presented in Table 6.2 contains four factors.

Table 6.2: ROTATED FACTOR MATRIX: VALUE SURVEY MODULE.

Variable	Description	Factor 1 Factor-score	Factor 2 Factor-score	Factor 3 Factor-score	Factor 4 Factor-score
Q25	Opportunity for advancement	0,81			
Q19	Co-operation	0,80			
Q22	High earnings	0,76			
Q21	Contribution to company	0,75			
Q24	Live in desirable area	0,74			
Q17	Security of employment	0,74			
Q27	Prestigious company	0,73			
Q15	Physical working conditions	0,72			
Q13	Challenging tasks	0,72		0,41	
Q29	Well-defined job situation	0,70			
Q28	Helping others	0,70			
Q18	Considerable freedom	0,67			
Q20	Consultation by superior	0,63			
Q12	Sufficient personal time	0,63			
Q26	Variety and adventure	0,58			
Q23	Serve one's country	0,50			
Q14	Little tension and stress	0,40			
Q33	Preference for large company		0,54		
Q35	Continuation of service		-0,44		
Q16	Good relationship with superior	0,42		-0,55	
Q37	Preferred manager				0,35

An analysis of the information in Table 6.2 shows that the factor matrix differed completely from the constructs and structure obtained by previous research (Hofstede, 1980b; Hofstede and Bond, 1984; Singh, 1990). The obtained factor matrix may be due to the prevalent ethnic differences in the South African society. The first one of the two extracted factors with eigenvalues greater than one may be termed work environment and the second work security.

As regards the Activism and Powerful Others-scale however, the factorial analysis done on the data (principal axis factoring with varimax rotation) yielded the same results as originally obtained by Levenson (1971). The factor analysis yielded three factors all of which have eigenvalues greater than one. The eigenvalues are presented in Table 6.3.

Table 6.3: EIGENVALUES: EXTRACTED FACTORS - ACTIVISM AND POWERFUL OTHERS-SCALE.

Factors	Eigenvalues	Percentage of variance	Cumulative Percentage
1	4,63	18,5	18,5
2	2,22	8,9	27,4
3	1,01	4,1	31,5

The information in Table 6.3 shows that the three extracted factors (eigenvalue > 1) declared 31,5% of the variance. The rotated factor matrix for these three factors is presented in Table 6.4.

Table 6.4: ROTATED FACTOR MATRIX: ACTIVISM AND POWERFUL OTHERS-SCALE.

Variable	Description	Factor 1 Factor-score	Factor 2 Factor-score	Factor 4 Factor-score
Q52	Good or bad fortune	0,74		
Q51	Luck	0,69		
Q60	Fate	0,62		
Q49	Accidental happenings	0,61		
Q50	Bad luck	0,48		
Q55	Luck determines events	0,46		
Q59	Important people	0,41		
Q58	Persons in authority		0,57	
Q43	Pleasing superiors		0,56	
Q41	Strong pressure groups		0,55	
Q42	Persons in control of authority		0,54	
Q40	Authoritative persons		0,49	
Q62	Desires of persons in authority		0,45	
Q53	Right time, right place		0,43	
Q54	Pre-determination		0,38	
Q45	Own actions			0,70
Q44	Own ability			0,69
Q47	Make plans work			0,61
Q48	Personal hard work			0,51
Q61	Own Leadership ability			0,48
Q56	Dexterity and skill			0,42

The information in Table 6.4 shows, that in general, the content of the questions classified under factor 1 relates to externality or chance factors. The questions classified under factor 2 relate to the dimension of powerful others. The questions classified under factor 3 relate to internality or the belief that one is in control of events in one's life. The information in Table 6.4 confirmed the construct validity of the Activism and Powerful Others-scale.

6.6 RELIABILITY

Reliability goes hand in hand with validity and involves the consistency or stability of a test score when the test is repeated or replicated. If a particular test, applied repeatedly to the same object, yields the same results each time, it is reliable (Smit, 1983, pp 28-29). Babbie (1989, p 56) defines reliability as "that quality of measurement method that suggests that the same data would have been collected each time in repeated observations of the same phenomenon". Mason et al (1989, p 420) view reliability as "the consistency or dependability of a test" and proceed to define reliability statistically as "the ratio of variance in the scores to variance in observed scores" (1989, p 266) and offer the formula

$$r_{xx} = \frac{T^2_t}{T^2_o} = \frac{T^2_t}{T^2_t + T^2_e}$$ where

r_{xx} = reliability
T^2_t = variance in true scores
T^2_o = variance in observed scores
T^2_e = variance of error.

6.6.1 COMPUTING RELIABILITY

Smit (1983) discerns three approaches to estimate reliability i.e. test-retest reliability, alternate forms reliability and internal consistency. The reliability estimate is determined by means of a correlation coefficient. The higher the numerical value of the obtained coefficient, the less the possibility of the effect of chance upon a test. The lower the obtained coefficient, the more the measure reflects chance factors (Mason et al, 1989, p 267).

6.6.1.1 TEST-RETEST RELIABILITY

Test-retest reliability boils down to two repeated administrations of the same test to the same group after a lapse of time. The two test scores obtained in this way are compared by means of correlational statistics. This procedure yields a reliability coefficient (r_{tt}) known as the coefficient of stability. The length of time between the two administrations may turn out to be a major problem. If the lapse of time is too short, carry-over effects like exercise and memory may effect the reliability. If the period is too long, maturation (biological, psychological and emotional processes that change subjects over time) may influence reliability (Smit, 1983, p 29; Dane, 1990, p 254).

6.6.1.2 ALTERNATE FORMS RELIABILITY

Alternate forms reliability involves comparing performances by the same group on two different but equivalent forms of the same test. Two equivalent forms of the test are administered to the same sample. A lapse of time between the two administrations is not necessary because two equivalent forms of the test are used (Smit, 1983, p 30). According to Smit (1983, p 30) the two equivalent forms must comply with certain requirements:

1. Both forms must be of equal length.
2. The same procedures for marking must apply to both forms.
3. Item homogeneity must be the same for both forms.
4. Items must be uniform in regard to content, representativeness and degree of difficulty.

If the time period between the two administrations is short, the reliability estimate is known as the coefficient of equivalence. If there is a long lapse of time, the reliability estimate is known as the coefficient of stability and equivalence (Smit, 1983, p 31).

6.6.1.3 INTERNAL CONSISTENCY

There are many methods for computing internal consistency, viz split-half reliability and the Kuder Richardson method, amongst other approaches. The split-half technique is one appropriate to assess the reliability of a questionnaire. It involves dividing the test into two equivalent halves and computing the correlation between the halves. A measure is usually divided by separating the odd and even numbered items (Smit, 1983, p 33). But this division of the test into two halves shortens the measure which in turn affects reliability. A correction to the reliability estimate has to be done to compensate for the shortened halves. Spearman-Brown advances the formula (Mason et al, 1989, p 268) to affect this correction

$$r_{tt} = \frac{2r_{oe}}{1 + r_{oe}}$$ where

r_{tt} = corrected reliability
r_{oe} = the reliability estimate of the split-half.

Guttman offers the following formula to affect the correction (Smit, 1983, p 24-35). This formula is independent of the requirements to calculate the correlation between the two halves.

$$r_{tt} = 2\left(1 - \frac{\sigma_A^2 + \sigma_B^2}{\sigma_t^2}\right)$$ where

σ_A^2 = variance of form A

σ_B^2 = variance of form B

σ_t^2 = variance of total group.

The Kuder-Richardson method, which usually yields higher reliability estimates because the measure is not split into two halves, is also employed to calculate internal consistency. The Kuder-Richardson formula 20 provides an estimate of the average split-half reliability without requiring actually splitting of the test (Smit, 1983, p 35). The Kuder-Richardson-formula 20 is

$$r_{xx} = \frac{k}{k - 1}\left(1 - \frac{\Sigma pq}{s_o^2}\right)$$

r_{xx} = reliability estimate
k = number of items on the test
p = the portion of people who respond correctly to each item
q = 1 - p
s_o^2 = Observed score variance
(Mason et al, 1989, p 269).

This Kuder-Richardson formula 20 is usually applied to obtain reliability coefficients when tests consist of dichotomously scored items. However, the Kuder-Richardson formula 20 may also be applied to tests comprising items which elicit more than two categories of response such as attitude scales. In the case of an item with more

than two response categories, the individual item variances are calculated and their sum substituted in the Kuder-Richardson formula 20 for

$$\sum_{i=1}^{n} p_i q_i$$

The Kuder-Richardson formula 20 used in the case of items which elicit more than two categories of response such as the case in hand the formula is (Ferguson, 1981, p 439):

$$\sum_{i=1}^{n} p_i q_i = \frac{n}{n-1}\left[1 - \frac{\sum_{i=1}^{n} s_i^2}{s^2}\right]$$

6.6.1.4 ITEM TOTAL RELIABILITY

Item total reliability is "an estimate of the consistency of one item with respect to other items on the measure" (Mason et al, 1989, p 256). Calculating an item total reliability involves correlating the score on one item with the total score on the rest of the items. The Kuder-Richardson formula 20 may be employed. A high correlation coefficient may be an indication of the entire instrument being reliable (Mason et al, 1989, p 256).

6.6.2 RELIABILITY OF THE VALUE SURVEY MODULE

Split-half and alpha reliability estimates were calculated by means of computer packages available on the main frame at the University of Pretoria. A split-half reliability estimate for unequal length of 0,79 was obtained. Because the partitioning of the questionnaire into two halves shortens the measure which in turn affects reliability, the Spearman-Brown correction to the

reliability estimate was done to compensate for the shortened halves (Mason et al, 1989, p 268; Smit, 1991, p 40). The Spearman-Brown correction yielded a reliability coefficient for unequal length of 0,88. An alpha coefficient of 0,90 was obtained for the Value Survey Module.

6.6.3 RELIABILITY OF THE ACTIVISM AND POWERFUL OTHERS-SCALE

Split-half and alpha reliability estimates were also calculated by means of computer packages available on the main frame at the University of Pretoria. A split-half reliability estimate for unequal length of 0,88 was obtained. A Spearman-Brown correction was also done to compensate for the shortened halves (Mason et al, 1989, p 268; Smit, 1991, p 40). The Spearman-Brown correction yielded a reliability coefficient for unequal length of 0,94. An alpha coefficient of 0,91 was obtained for the Activism and Powerful Others-scale.

6.7 SUMMARY

In this chapter the psychological tests used in the study were discussed. Attention was given to the construction and development of both Hofstede's Value Survey Module and Levenson's Activism and Powerful Others-scale. The different approaches in determining validity and reliability estimates were discussed in some detail and split-half and alpha reliability estimates calculated for the Value Survey Module and the Activism and Powerful Others-scale.