

## **CHAPTER 5**

### **EMPIRICAL STUDY**

With reference to the research methodology, as discussed in chapter 1, section 1.7.2, phase 2, steps 1-4, the aim of this chapter is to discuss the empirical study of the research. The chapter begins with a discussion on the selection of the population and sample for the research. Secondly, the compilation and motivation for the psychometric instruments will be reviewed. This is followed by the discussion of the procedures and statistical techniques used in the empirical study. The chapter ends with the formulation of the central hypothesis.

#### **5.1 THE POPULATION AND SAMPLE**

The population researched in this study were South African community service doctors in full time service from KwaZulu-Natal state hospitals in South Africa ( $N = 226$ ).

From this population, the sample drawn consisted of 41 community service doctors ( $N = 41$ ). This sample represented 18,14% of the total number of community service doctors in KwaZulu-Natal state hospitals.

In this research, an accidental sample was drawn, by convenience. According to Kerlinger (1986, p. 120), an accidental sample refers to the available sample at hand, which, although it is the weakest form of sample, is most probably the most frequently used. All participants in the sample were in their post internship year and were obliged to complete compulsory community service, which was introduced for the first time in South Africa.

Written consent was requested from the superintendents of the hospitals in the KwaZulu-Natal area, requesting the participation of the community service doctors in the research.

## **5.2 THE MEASURING INSTRUMENTS**

The research methodology (sec. 1.7.2, step 2, phase 2) of the empirical study, namely the discussion and justification of the psychometric instruments, is addressed.

The measuring battery used for the purpose of this study includes seven quantitative instruments, one of which measures stress, another burnout, four measure salutogenic constructs and the remaining, biographical data. Each instrument is discussed in terms of its aim, development, rationale, description, administration, interpretation, reliability, validity and justification for inclusion in this research.

### **5.2.1 Stress Diagnostic Survey (SDS)**

The SDS is used to measure the theoretical concept of job related stress.

#### *5.2.1.1 Development*

Stress Research Systems developed the SDS with a view to understanding stress in the work environment. It is described by Ivancevich and Mateson (1980).

#### *5.2.1.2 Rationale*

The rationale of the SDS is that the higher the respondent's total score (the total of the extent to which various individual level stressors are sources of stress), the higher the levels of job stress experienced by that individual (Arumugam, 1992, p. 54).

#### *5.2.1.3 Description*

The SDS is a questionnaire, comprising 30 items rated on a 7 point scale. A score of 1 describes a work condition never being a source of stress while a score of 7 describes a work condition always being a source of stress. The questionnaire

measures 5 individual level stressor dimensions (Ivancevich and Matteson, 1980, Arumugam, 1992):

- Role conflict is present whenever compliance with one set of demands makes compliance with another set of demands difficult, objectionable or impossible.
- Role ambiguity is a lack of clarity about one's role, job objectives, the scope of responsibilities, or one's job.
- Work overload can be divided into two categories: quantitative overload and qualitative overload. Quantitative overload occurs when there is too much to do in too limited a period and qualitative overload which refers to a state of being in which work demands exceed capability.
- Responsibility for people opposed to things.
- Career development stress includes those aspects which influence the individual's interaction with the organisational environment, which, in turn, influences that person's perception of the quality of his or her career progress.

#### 5.2.1.4 Administration

The SDS can be administered individually or within a group. It is essentially a self-administered questionnaire, where respondents follow a set of simple instructions. Next to each of the statements, the respondent is requested to mark the number (1-7, where 1 represents never and 7 represents always) which best describes how frequently the condition described is a source of stress for him or her (Ivancevich & Matteson, 1980). The respondent must give only one spontaneous answer to each question and is encouraged not to omit any items. On average, it takes 15 to 20 minutes to complete, but there is no time limit.

#### 5.2.1.5 *Scoring*

The SDS is scored using a scoring key (Arumugam, 1992). Role conflict is scored by adding the respondent's score for item numbers: 1, 7, 13, 19 and 25. Role ambiguity is scored by adding item numbers: 2, 8, 14, 20 and 26. Work overload quantitative is scored by adding item numbers: 3, 9, 15, 21 and 27. Work overload qualitative is scored by adding item numbers: 4, 10, 16, 22 and 28. Career development is scored by adding item numbers: 5, 11, 17, 23 and 29. Responsibility for people is scored by adding item numbers: 6, 12, 18, 24 and 30. All the six subtotals can be added to attain a total stress score.

#### 5.2.1.6 *Interpretation*

Each of the five individual level stressor categories can be classified into low, moderate or high stress according to the sum of their item scores. For any stressor category, low stress is less than 10, moderate stress is between 10 and 24 and high stress is greater than 24. The total stress score can be attained by adding all the subcategories. The total score can also be classified on the same low-moderate-high scale (Ivancevich & Matteson, 1980).

#### 5.2.1.7 *Reliability*

(No extensive reliability and validity studies were available for the SDS.)

An internal reliability coefficient of 0,82 was reported by Boyd (1982) in his analysis of the relationship between stress levels of the SDS and type A/B behaviour on the Jenkins Activity Survey. In another study, Arumugam (1992, p. 61) found significant association between levels of experienced stress and the stressor categories of the SDS in a sample of Indian Pentecostal Ministers.

#### 5.2.1.8 *Validity*

No validity studies were available regarding the use of the SDS (Arumugam, 1992:118). However, the SDS is recommended for its face validity.

#### *5.2.1.9 Justification*

As documented in the literature (see chap. 2), frequently cited sources of stress are role demands, work overload, responsibility of caring and career development concerns (Lee & Ashforth, 1990, p. 743). The SDS has attempted to measure these sources of stress and other well known sources of stress which are reported to be antecedents in the burnout process. Despite the limited research studies available, this instrument has also proved to be a reliable, brief, easily interpretable and adequate measure for the assessment of work stress levels. For these reasons, the SDS was chosen to assess the levels of stress in the sample of community service doctors.

### **5.2.2 The Maslach Burnout Inventory (MBI)**

The MBI is used to measure the theoretical concept of burnout.

#### *5.2.2.1 Development*

In an attempt to overcome confusion about the nature and measurement of the burnout concept, Maslach (1973, in Schaufeli et al, 1993), began reviewing the diverse research findings on the burnout concept and was able to extract general themes that emerged. This attempt resulted in the development of a standardised scale measure, the Maslach Burnout Inventory (MBI), by Maslach and Jackson (1981).

The MBI was made up of three subscales which were developed from a factor analytic study involving 2118 subjects from various people-oriented vocations (845 social security administration public contact employees, 142 police officers, 231 nurses, 125 agency administration, 222 teachers, 97 counsellors, 91 social workers, 68 probation officers, 63 mental health workers, 86 physicians, 40 psychologists and psychiatrists, 31 attorneys and 77 others) (Maslach & Jackson, 1981).

As stated by Maslach (1982b, p. 38), “the use of the MBI by many researchers has facilitated the comparison and integration of various research findings, is generating a data bank for standardized norms and is stimulating new ideas about patterns or phases of burnout”.

#### 5.2.2.2 *Rationale*

The MBI contains three subscales that assess the different aspects of experienced burnout. Maslach and Jackson (1981, p. 1) describe these as follows:

- The emotional exhaustion subscale assesses feelings of being emotionally overextended and exhausted by one’s work.
- The depersonalisation subscale measures an unfeeling and impersonal response towards recipients of one’s service, care, treatment or instruction.
- The personal accomplishment subscale assesses feelings of competence and successful achievement in one’s work with people.

Each subscale has two dimensions: frequency (how often people have these feelings) and intensity (the strength of these feelings). Six scores are calculated for each respondent: emotional exhaustion frequency and intensity, depersonalisation frequency and intensity, and personal accomplishment frequency and intensity (Maslach & Jackson, 1981).

Burnout is conceptualised as a continuous variable ranging from low to moderate to high degrees of experienced feeling. It is not viewed as a dichotomous variable which is either present or absent (Maslach & Jackson, 1986). A high degree of burnout is reflected in high scores on the emotional exhaustion and depersonalisation subscale, and low scores on the personal accomplishment subscale. A moderate degree of burnout is reflected in moderate scores on the three subscales. A low degree of burnout is reflected in low scores on the emotional exhaustion and depersonalisation subscale, and in high scores in the

personal accomplishment subscale. Scores are considered high if they are in the upper third of the normative distribution, moderate if they are in the middle third and low if they are in the lower third (Maslach & Jackson, 1981).

#### 5.2.2.3 *Description*

As described in the rationale, items for the MBI were designed to measure hypothetical aspects of the burnout syndrome. There are 22 items of job-related feelings in a two-scale format (Maslach and Jackson, 1981). The nine items in the emotional exhaustion subscale describe feelings of being emotionally overextended and exhausted by one's work. Five items in the depersonalisation subscale describe an impersonal response towards recipients of one's care or service. The subscale of personal accomplishment contains eight items that describe feelings of competence and successful achievement in one's work with people. Each statement is rated on two dimensions: frequency and intensity. The frequency scale is labelled at each point and ranges from 0 ("never") to 7 ("every day"). The intensity scale ranges from 0 ("never") to 7 ("major, very strong").

In order to avoid sensitisation and response bias, the MBI is introduced to the respondent as the Human Services Survey, a questionnaire comprising a list of statements about job-related attitudes (Maslach & Jackson, 1981).

#### 5.2.2.4 *Administration*

The MBI is a self-administered questionnaire and takes about 20 to 30 minutes to fill in. The respondent is asked to answer 22 statements regarding their job-related feelings on a scale of 0 to 7, where 0 represents never having this feeling about one's job and 7 represents having this feeling everyday. For each question, the respondent is also asked to express how strong the feeling is when experienced where 0 represents never, 1 represents mild and 7 represents a major or very strong feeling (Maslach & Jackson, 1981).

#### 5.2.2.5 *Scoring*

According to Maslach and Jackson (1981, p. 3), each respondent's test form is scored by use of a scoring key, which contains directions for scoring each subscale and dimension.

#### 5.2.2.6 *Interpretation*

According to Maslach and Jackson (1981), the scores derived from the scoring key can then be coded as low, moderate or high by using the numerical cut-off points listed on the scoring key. Means and standard deviations for each subscale are computed and can be compared to the normative data.

Given the considerable factor-analytic support for their separation and limited knowledge about the relationship between the three subscales of burnout, the scores for each subscale are considered separately (Maslach & Jackson, 1981; 1986). For the purpose of this research, the frequency and intensity subscales were analysed independently; recent burnout studies, however, report relatively high correlations between the frequency and intensity subscales of the MBI (Maslach, Jackson & Leiter, 1996).

Those respondents who score high on the emotional exhaustion subscale, are predicted to be experiencing high levels of psychological and physiological strain (Leiter, 1992; Perlman & Hartman, 1981).

Those respondents who score high on the depersonalisation subscale, portray defensive behaviour intended to avoid an unwanted demand. Here the individual attempts to staunch the depletion of emotional energy by treating others as objects or numbers rather than as people (Ashforth & Lee, 1990, p. 745; Kahil, 1988; Maslach, 1982).

Those who score high on personal accomplishment are predicted to be well adjusted and in control of the demands in their environment. Personal

accomplishment has been reported to be positively associated with self-efficacy and negatively linked to a sense of helplessness (Rosenbaum, 1988).

#### 5.2.2.7 Reliability

According to Maslach and Jackson (1981), the internal consistencies as estimated by the cronbach alpha coefficients are as follows (n=1316):

*Emotional exhaustion: 0,90 (frequency) and 0,87 (intensity).*

*Depersonalisation: 0,79 (frequency) and 0,76 (intensity).*

*Personal accomplishment: 0,71 (frequency) and 0,73 (intensity).*

Data on test-retest reliability after a period of two to four weeks for the subscales ranged from 0,53 to 0,82. Although these coefficients range from low to moderately high, all are significant beyond the 0,001 level (Maslach & Jackson, 1981).

In another study, Maslach and Jackson (1986), and Lahoz and Mason (1989) in Rothmann and Fourie (2002), reported cronbach alpha coefficients varying from 0,71 to 0,90 for the three subscales of the MBI; test-retest reliability varied from 0,60 to 0,82 and 0,54 to 0,60 (applied after 1 year).

#### 5.2.2.8 Validity

The convergent validity of the MBI was demonstrated in several ways, namely, with peer ratings, job dimensions associated with burnout and stress outcomes (Maslach & Jackson, 1981). Numerous studies report that the MBI has significant concurrent and predictive validity (Rafferty et al, 1986). External validation of the MBI comes from analyses of its convergence with peer ratings, job dimensions associated with burnout and stress outcomes (Maslach & Jackson, 1984 in Rothmann, Malan & Rothmann, 2002).

According to Maslach and Jackson (1981), discriminant validity of the MBI has been established by distinguishing it from other psychological constructs, such as

job dissatisfaction, that might be presumed to be confounded with burnout. Also, none of the burnout subscales were found to be significantly correlated with social desirability (Maslach & Jackson, 1981).

Finally, Van Dierendock and Schaufeli (1994) highlight the cross-cultural validity of the MBI.

#### *5.2.2.9 Justification*

The literature confirms that the MBI is an adequate and interpretable measure for the purpose of assessing the level of burnout in professionals who do “people work”(Van Dierendock & Schaufeli, 1994). Also, as mentioned above, the reliability and validity of the instrument has been established. Finally, the replicable, three factor structure of burnout conceptualised by Maslach (1982a), has been confirmed in majority of studies and is becoming widely accepted as the most comprehensive conceptualisation of burnout. For this reason, the MBI was chosen to determine the levels of burnout in the community service officers in KwaZulu-Natal hospitals.

### **5.2.3 Orientation to Life Questionnaire (OLQ)**

The OLQ (Antonovsky, 1987b, p. 77) is used to measure an individual’s sense of coherence.

#### *5.2.3.1 Development*

Antonovsky (1987b, p. 163), states that evidence does lend itself to indicate that a relationship does exist between the sense of coherence (SOC) and health. The OLQ was, therefore, constructed to measure the SOC construct in order to test the hypothesis that the SOC is causally related to health status (Antonovsky, 1987b, p. 63). The twenty-nine item SOC Hebrew questionnaire was tested for the first time in 1983 on an Israeli national sample. The final form of the questionnaire appeared in 1985 and consisted of 29 items; eleven measure comprehensibility, ten measure manageability and eight measure meaningfulness.

This questionnaire has its basis in the salutogenic as opposed to the pathogenic-oriented approach. In the salutogenic model, (Antonovsky, 1987 in Wolf & Ratner, 1999, p. 182) there is a suggestion that a SOC is the key determinant in the maintenance of health and prevention of breakdown, and that individuals with a strong SOC have the ability to perceive stressors as manageable, meaningful and comprehensible.

Antonovsky (1990) developed the OLQ by using experimental subjects who had experienced serious and irreversible trauma in their lives; these people were characterised by others as functioning remarkably well despite the trauma they had experienced. The sample consisted of 30 males and 21 women of which 47 ranged between 21 and 91 in age, and of which four were teenagers.

The subjects were categorised according to a seven-point scale ranging from weak to strong sense of coherence. The formulation of the above led Antonovsky (1987b, p. 77) to the development of the technique of “mapping sentence for questionnaire design” which was utilised to determine which items would be included in the questionnaire.

#### *5.2.3.2 Rationale*

The SOC is a major determinant of the location and movement of an individual's state of wellness along the health-disease continuum (Antonovsky, 1984; 1987a; 1987b; 1993b; 1996; Wolf & Ratner, 1999, p. 183). The stronger an individual's SOC is, the greater the chance he or she is likely to maintain his or her location or even improve it. One's score on the SOC determines one's ability to cope with stressful situations. Antonovsky and Sagy (1986, p. 216) advocate that individuals with a high SOC are less likely to have a tendency to perceive situations as ego threatening and anxiety arousing. In accordance with this, the higher the score on the OLQ, the stronger the SOC of the respondent will be.

### 5.2.3.3 *Description*

The OLQ consists of 29 five-facet items rated on a seven-point, bipolar graphic rating scale (Antonovsky 1993b, p. 726). It comprises 11 items measuring comprehensibility, 10 manageability, and 8 meaningfulness. Respondents indicate the extent to which they agree or disagree with the items (Antonovsky, 1990). Overall, the OLQ is a self-report instrument aimed at evaluating an individual's tendency to successfully apply coping mechanisms (Antonovsky, 1984, p. 117).

### 5.2.3.4 *Administration*

The OLQ can be administered individually or within a group. It is essentially a self-administered questionnaire, whereby respondents follow a set of simple instruction. Following a seven-point Likert scale, the respondent then marks the number which best expresses the extent to which the statements are applicable to him or her. The respondent must give only one spontaneous answer to each question and is encouraged not to omit any items. On average, it takes 15 to 20 minutes to complete, but there is no time limit.

### 5.2.3.5 *Scoring*

The instrument is scored by numerically adding all 29 items. The following steps are employed in the scoring of this instrument:

Step 1: Items 1, 4, 5, 6, 7, 11, 13, 14, 16, 20, 23, 25, and 27 are formulated negatively and, therefore, have to be reversed before being added, so that the high scores always represent a strong sense of coherence.

Step 2: The values of the following items are added for the total scores on

Comprehensibility (c) = 1, 3, 5, 10, 12, 15, 17, 19, 21, 24 and 26

Manageability (ma) = 2, 6, 9, 13, 18, 20, 23, 25, 27 and 29

Meaningfulness (me) = 4, 7, 8, 11, 14, 16, 22 and 28

#### 5.2.3.6 *Interpretation*

The theoretical range on the OLQ is 29 to 203. A high score on the OLQ represents a strong sense of coherence, whereas a low score indicates a respondent low in sense of coherence. Antonovsky (1984) cautions that having a strong sense of coherence does not mean that a respondent views his or her entire world as comprehensible, manageable and meaningful; rather it should be interpreted as being on the healthy side of the health ease/dis-ease continuum. Normative data on the SOC present means and standard deviations for a variety of samples, against which comparisons can be made. For a group of Israeli army officers – a group expected to have very high SOC - the mean was 160,44 and standard deviation 16,69 compared to a sample of US undergraduates who had a mean of 129,5 with a standard deviation of 24,5 (Antonovsky, 1984, p. 119).

#### 5.2.3.7 *Reliability*

The consistently high level of cronbach's alpha as indicated in 26 studies, ranging from 0,84 to 0,93, indicates a high degree of internal consistency and reliability (Antonovsky, 1987b, p. 82; 1990). Antonovsky (1987b, p. 82), reports that this high internal consistency and reliability has been found in a variety of populations, with different languages and cultures. Rothmann (2000) in Rothmann & Basson (2002), reported an alpha coefficient of 0,89 for the OLQ which may be regarded as acceptable.

Test-retest reliability varies between 0,91 for a six week period, 0,76 after one year and 0,41 for a two year period in a sample of Israeli medical students (Carmel & Bernstein, 1990). Fiorentino (1986) reports a one year test–retest correlation of 0,78. Antonovsky and Sagy (1986) report a test-retest correlation of 0,54 after one year and 0,55 after two years.

#### 5.2.3.8 *Validity*

Using the facet method of constructing scales, an item was only included after four colleagues, familiar with the theory, had independently concurred that it indeed referred to only one of the three SOC components - a form of face and content validity (Antonovsky, 1993b, p. 727; 1987b, pp. 82-86). Flannery et al (1994 cited in Wolf & Ratner, 1999, p. 190) demonstrate the construct validity of the scale via principal component analysis where three factors emerged as hypothesised by Antonovsky.

Overall, the majority of the studies on the OLQ support the validity of the scale. The following is an example of some available publications:

Rumbaut (1983 in Antonovsky, 1987b, p. 83) and Dana et al (1985 in Antonovsky, 1987b, p. 83), in a criterion-related validity study (concurrent validity) found validity coefficients of 0,64 and 0,72 between the OLQ and Rumbaut's 22-item SOC scale.

With regard to construct validity of the OLQ, it was found that there is a negative relationship between the OLQ and experienced stress, and that the OLQ correlates negatively with the "State – Trait Anxiety Inventory – Trait" and the "Beck Depression Inventory" (Frenz et al, 1993 in Rothmann & Basson, 2002). Payne (1982, pp. 2-3) reports that the SOC score was consistently and significantly related to all positive health measures, while being significantly and negatively related to all illness measures. This was verified by numerous other studies: (Antonovsky & Sagy 1986; Carmel & Bernstein 1990; Radmacher & Sheridan, 1989 in Carstens & Spanenberg, 1997, p. 1215; Fiorention, 1986; Geyer, 1997; Langius, Bjorvell & Antonovsky, 1992, p. 169).

Antonovsky (1993b, pp. 727-730) concludes the reliability and validity of the OLQ as follows: "The cronbach's alpha coefficient reviewed in 29 studies is between 0,91 and 0,85 – the highest consistency found for different population groups in different cultures and language groups in the West." Antonovsky (1993b) also

reports that the entire spectrum of focus on the test-retest reliability produced a reliability coefficient between 0,41 and 0,97.

From the above evidence, it can be concluded that the 29-item OLQ can be regarded as a valid instrument that measures what it claims to measure.

#### *5.2.3.9 Justification*

On a conceptual level, the sense of coherence is the central construct in the salutogenic model – developed by Antonovsky (1987). The OLQ was found to best measure the SOC construct, in that this questionnaire operationalised the concept in terms of its definition (see chap. four).

One of the key strengths of the OLQ, is that it is universally meaningful, and can be utilised across gender, social class, religion and culture (Antonovsky 1993b, p. 726). Furthermore, it has normative data from around the world and from a variety of occupational groups – implying it could be used for the sample of community service doctors in this research.

In addition, the OLQ was found to be valid and reliable – bearing sound psychometric properties. Finally, Antonovsky (1993b, p. 726), advances that the SOC is based on good theory, is a sound instrument and was empirically tested on a large scale.

### **5.2.4 Personal Views Survey (PVS)**

The PVS is used to measure the theoretical concept of hardiness.

#### *5.2.4.1 Development*

The PVS is an adapted version of the Hardiness Questionnaire as developed by Kobasa (1979). Shortly after research showed that psychological stress was associated with illness, researchers began to search for psychosocial characteristics that might moderate the stress-illness relation (Dohrenwend &

Dohrenwend, 1974 in Funk, 1992, p. 33). During this period, the concept of hardiness was introduced by Kobasa (1979). In the early years of hardiness research, 19 different instruments were developed to measure this construct (Funk, 1992, p. 336).

Many of these scales did not distinguish between high-stress/high-illness and high-stress/low-illness subjects and, therefore were excluded as adequate measurement of the hardiness construct. The development of the PVS went through a stage development process (Wilmans, 1995, p. 73). According to Funk (1992), it started off as having six scales and this led to the development of the 71-item Unabridged Hardiness Scales (UHS). The UHS was later replaced by the creation of the two short forms: the 20-item Abridged Hardiness Scale (AHS) and the 36-item Revised Hardiness Scale (RHS). More recently, as a result of problems with these scales, the 50-item Personal Views Survey (PVS) was developed by the Hardiness Institute (Funk, 1992, p. 336).

#### *5.2.4.2 Rationale*

The concept of hardiness falls under the umbrella of salutogenesis and corresponds to other constructs like the sense of coherence, learned resourcefulness and self efficacy (Strümpfer, 1990). Kobasa (1979), who was instrumental in developing the hardiness concept, proposed that people who experienced high levels of stress, yet remained healthy had a different personality structure from people who experienced high levels of stress and became ill. In keeping with the salutogenic line of thinking, Manning, Williams and Wolfe (1988, p. 205) advocate that hardiness is a constellation of personality characteristics, namely, commitment, control and challenge that function as a resistance resource in the encounter with stressful life events. The rationale of the PVS is that, the higher the respondent's score on the three components, the stronger the hardiness of that individual.

#### 5.2.4.3 *Description*

The questionnaire consists of 50 self-descriptive items (Kobasa, 1979b). Fifteen of the items measure commitment, 17 measure control and 18 measure challenge. The respondents choose on a four-point Likert scale the extent to which he or she agrees with each statement. A score of zero indicates total disagreement with the statement; a score of three indicates high agreement with the statement. The items are worded strongly, encouraging easier decision-making in selection.

#### 5.2.4.4 *Administration*

The PVS can be answered individually or in groups. The respondent is provided with a self-completion instrument and is asked to allocate a rating on a scale ranging from zero (not at all) to three (completely true) to indicate the extent to which he or she agrees with each statement (Kobasa, 1979b). The respondents are requested to answer all statements on the basis of the way they presently feel. There is no time limit, but respondents are encouraged to indicate their first spontaneous response.

#### 5.2.4.5 *Scoring*

The PVS involves reversing scores and then adding the scores (Kobasa, 1979b). The PVS can be scored in the following manner:

Step 1: Reverse eleven of the items 1, 2, 3, 4, 5, 22, 23, 24, 25, 26 and 27, so that the high score always indicates a strong hardiness.

Step 2: The value of the following items are added for the score on commitment: 1, 8, 11, 14, 17, 20, 23, 26, 29, 32, 38, 39, 44, 47 and 50.

Step 3: The values of the following items are added for the score on control: 3, 4, 5, 7, 10, 13, 16, 19, 22, 25, 28, 31, 34, 35, 42, 45 and 48.

Step 4: The values of the following items are added for the score on challenge: 2, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 37, 40, 41, 43, 46 and 49.

#### 5.2.4.6 *Interpretation*

A high score indicates a respondent high in hardiness whereas a low score indicates a respondent low in hardiness. Kobasa (1979b, p. 3), and other behaviour control theorists, such as, Lefcourt, (1966), Antonovsky, (1979), Lazarus (1966), and Maddi (1990) advance that hardy persons are considered to possess the following three general characteristics:

- the belief that they can control or influence the events of their experience
- an ability to feel deeply involved in or committed to the activities of their lives
- the anticipation of change as an exciting challenge to further development

One's level of hardiness is a predictor of the extent to which the stress and illness relationship will be buffered (Kobasa, 1982).

#### 5.2.4.7 *Reliability*

Kobasa (1979a, pp. 418-419) reports correlation coefficients of 0,85 for commitment, 0,68 for control and 0,70 for challenge – this is consistent with several other research studies done on the hardiness construct. Overall, hardiness as a construct, has been found to significantly moderate the stress-illness relationship (Kobasa et al, 1982, p. 175).

#### 5.2.4.8 *Validity*

The internal validity of the subscales of the PVS show a significant internal validity of 0,85 for commitment, 0,70 for control and 0,71 for challenge; the combination for the subscales indicate a correlation of 0,61 over a period of five years (Kobasa 1982). According to Kravetz, Drory and Florian (1993, p. 237), the PVS is reported to have adequate internal consistency and test-retest stability. In this

study, the cronbach's alpha was approximately 0,70 for each of the three subscales. The correlation between two administrations of the full scale within a two week interval was 0,60. Also, according to Maddi and Kobasa (1984, p. 267), this test shows promising construct validity.

#### *5.2.4.9 Justification*

The PVS was chosen for this study because it was found to adequately operationalise the concept of hardiness as discussed in chapter four.

Two of the empirical aims of this research were to report on the relationships between stress, burnout and salutogenic functioning, and to ascertain whether salutogenic construct scores have the ability to differentiate between copers and noncopers. This was thus considered an ideal test in meeting these empirical objectives because, on a conceptual level, it is reported that hardiness acts as a predictor of the extent to which stress and illness will be buffered by hardiness Kobasa (1979b, p. 3).

The evidence on the reliability and validity of the instrument also made it a viable option to be used in the study.

### **5.2.5 Internal-External Control (I-E scale)**

The I-E scale is used for measuring the theoretical concept of locus of control.

#### *5.2.5.1 Development*

According to Lefcourt (1966, p. 206), under various rubrics and from diverse orientations, investigators have concerned themselves repeatedly with man's ability to control his personal environment. Concepts, such as competence, helplessness and hopelessness are some concepts that have been used to describe the degree to which an individual is able to control the important events occurring in his or her life.

The first attempt to measure individual differences in a generalised expectancy of belief in external control as a psychological variable was carried out by Phares in 1957 in his study of chance and skill effects on expectancies for reinforcement (Rotter 1966, p. 9; 1989, p. 489; 1992). This gave rise to development of a Likert-type scale with 13 items stated as internal attitudes and 13 items stated as external attitude. This initial scale was subsequently tested and changed several times - finally culminating in the Rotter's 29 item I-E scale.

#### *5.2.5.2 Rationale*

Rotter (1966, p. 10) advocates that the items in the I-E scale deal exclusively with the subject's belief about the nature of the world. This instrument is the concern with the respondent's expectations about how reinforcement is controlled. The rationale behind this is the extent of the individual's belief that he or she is in control of life events (internal locus of control), or that events are controlled by such external sources as chance and powerful others (external locus of control) (Kravetz et al, 1993). Hence, the test is considered to be a measure of a generalised expectancy. Such generalised expectancy may correlate with the value the subject places on control (internal or external). High scores reflect internal locus of control.

#### *5.2.5.3 Description*

This instrument is made up of 23 locus of control items and six filler items in forced-choice format (Hurrell & Murphy, 1991, p. 135). The aim of filler items is to reduce the transparency of the aim of the questionnaire. This 29-item questionnaire includes two statements for each item (Kravetz et al, 1993). However, none of the items is directly addressed to the preference for internal or external control.

#### *5.2.5.4 Administration*

The I-E scale can be administered individually or in groups. It is a self-completing instrument and respondents are requested to select one statement of a pair of

statements which he or she believes to be more true of themselves (Rotter, 1966). There are no right or wrong answers and respondents are asked to answer spontaneously. There is no time restriction.

#### 5.2.5.5 *Scoring*

The I-E scale is an additive one and the items are not comparable Rotter (1966, p. 11).

The following steps guide the scoring of the instrument:

Step 1: Ignore the six filler items: 1, 8, 14, 19, 24 and 27.

Step 2: To calculate external control, the following items must be added: 2a, 3b, 4b, 5b, 6a, 7a, 9a, 10b, 11b, 12b, 13b, 15b, 16a, 17a, 18a, 20a, 21a, 22b, 23a, 25a, 26b, 28b and 29a.

Step 3: To calculate the internal control, subtract the total score in step 2 from 23. The result is the internal control score.

#### 5.2.5.6 *Interpretation*

A review of studies by Rotter (1966, pp. 19-26) and Lefcourt (1966), indicate a strong relationship between internal control and

- attempts to control the environment
- the need for achievement motivation
- the ability to resist subtle influences made upon them

According to Hurrell and Murphy (1991, p. 137) one of the most consistent and widespread findings in the locus of control literature is the association between belief in external control and self-reports of ill health.

#### 5.2.5.7 *Reliability*

Rotter (1966, p. 25) states that overall, studies represent an unusually consistent set of findings even across populations. A moderate to high correlation coefficient of 0,49 to 0,83 was discovered. Kravetz et al (1993) found cronbach's alpha of 0,72 for the Hebrew version of the I-E scale.

Rotter (1989, p. 491) explains that in studying the I-E scale, the developers did not try for a high alpha (ie, the accepted measure of internal consistency), because they assumed that the correlations among different behavioural referents for the concept were positive but low and that it was possible to get very high alphas by asking the same question over and over again; but the predictive limitations of such tests are obvious.

Item analysis and factor analysis indicate sufficiently high internal consistency compared to other additive scales. Test-retest reliability has proved to be satisfactory and the scale was found to correlate with other methods for assessing the same variable such as the questionnaire (Rotter, 1966).

#### 5.2.5.8 *Validity*

With regard to construct validity of the I-E scale, significant evidence comes from the predicted differences in behaviour for individuals above or below the median of the scale or from correlations with behavioural criteria and similar scales such as the Liverant's 60-item scale Rotter (1966, p. 25). Furthermore, it is reported that there is consistency.

Rotter (1966, p. 24), indicates that differences in means of selecting populations are generally a weak criterion of validity. If taken into consideration that the differences obtained for different types of populations are consistent, this is an indication of strong validity (Marais, 1997).

#### 5.2.5.9 *Justification*

“Locus of control is currently one of the most studied variables in psychology, however relatively few locus of control studies have been conducted in work settings, and only a limited number of these are concerned with work demands and their health-related outcomes” Hurrell and Murphy (1991). By including the locus of control construct in the study, it was an opportunity to contribute to the body of knowledge of the construct in work settings - which is currently lacking.

Internal locus of control is linked to coping effectively with environmental demands. In an attempt to assess if salutogenic constructs could differentiate between copers and noncopers, this was considered an ideal construct to assist in answering this question.

The reliability and validity reports of the instrument further supported its use in this study.

#### **5.2.6 Self-Control Schedule (SCS)**

The SCS is used to measure the theoretical concept of learned resourcefulness.

##### 5.2.6.1 *Development*

Rosenbaum (1980, p. 109) developed the self-report instrument to assess individual tendencies to apply self-control methods to the solution of behavioural problems. During this period, there was heightened interest in the nature of self-management and the application of self-management techniques in behaviour therapy. Rosenbaum was particularly interested in the role of learning histories and individual differences in self control behaviour

#### 5.2.6.2 *Rationale*

The SCS is aimed at assessing individual's tendencies to apply self-control methods to the solution of behavioural problems (Rosenbaum 1989, p. 253). As reported by Rosenbaum and Palmon (1984, p. 246), self-control behaviours were categorised into four content areas which parallel the three conceptual dimensions of resourcefulness. These are the use of cognition and self-instruction which images self-control, the application of problem-solving strategies reflecting self-direction and the belief in one's coping effectiveness represents self-efficacy.

According to Rosenbaum (1980, p. 110), self-controlling responses were assumed to be cued by any internal event such as anxiety, pain or thought that disrupts the effective performance of a target behaviour. Self-controlling responses are then directed at reducing the interference of such events. Essentially then, the rationale of the SCS enables the definition of a) the conditions in which self-controlling responses are likely to occur and b) the function of self-control behaviour. A high score on the SCS indicates high learned resourcefulness and stronger self-control behaviour (Rosenbaum, 1980).

#### 5.2.6.3 *Description*

As described in the rationale, the SCS was designed to measure the four content areas described in the previous section, namely self-statements, problem-solving strategies, delaying immediate gratification and self-efficacy. The SCS comprises 36 items of which 12 refer to control of emotional and psychological sensations, 11 to the respondent's tendencies to employ problem-solving strategies, 4 to the respondents perceived ability to delay immediate gratification and 9 to indicate general expectations for self-efficacy Rosenbaum (1980, p. 114).

#### 5.2.6.4 *Administration*

The SCS can be answered individually or in groups. The SCS is a self-report instrument rated on a six-point scales ranging from +3 to -3. A score of +3 indicates behaviour that is very characteristic of the respondent and a score of -3

indicates behaviour that is very nondescriptive of the respondent, (Rosenbaum & Palmon, 1984, p. 246). There is no time limit for completing the questionnaire and only one answer may be given to each question.

#### 5.2.6.5 *Scoring*

The SCS is an additive one and the total score is the sum of the individual items; it indicates the respondent's learned resourcefulness (Rosenbaum, 1980). The following steps guide the scoring of this instrument:

Step 1: Eleven of the items are reversed so that the high score always represent strong self-control behaviour: 4, 6, 8, 9, 14, 16, 18, 19, 21, 29 and 35.

Step 2: Add all scores. The sum total is the total score on self-control behaviour.

#### 5.2.6.6 *Interpretation*

According to Rosenbaum (1980, p. 110), an individual who measures strongly on the SCS will employ the following coping strategies in dealing with stressful life events will

- use cognitions and “self-statements” to control emotional and physiological responses
- apply problem solving strategies
- have the ability to delay immediate gratification
- possess perceived self-efficacy

A high score indicates a respondent high in learned resourcefulness, with a tendency to apply self-control to the solution of common behavioural problems; a low score indicates a respondent low in learned resourcefulness (Rosenbaum, 1980).

#### 5.2.6.7 *Reliability*

The alpha coefficients for the SCS obtained ranged from 0,78 to 0,91 (Rosenbaum, 1988).

In a study conducted by Rosenbaum and Palmon (1984, p. 246) involving 600 subjects, the test-retest reliability over a four week testing period was 0,96 – indicating fairly high stability of test scores. Rosenbaum (1980) found a test-retest reliability of 0,86 for the SCS (Clanton, Rude & Taylor, 1992, p. 134)

The SCS is noted for high internal consistency amongst items with alpha coefficients ranging from 0,78 to 0,86 and is considered satisfactory for experimental purposes (Rosenbaum, 1980, p. 115). Internal consistency reliabilities ranged from 0,78 to 0,84 for the Hebrew version of the SCS (Clanton et al, 1992).

According to Rosenbaum and Palmon (1984, p. 246), a study conducted on a sample of approximately 1000 undergraduates using the SCS - by Redden, Tucker and Young in the US - revealed an alpha reliability of 0,82.

#### 5.2.6.8 *Validity*

When compared to other behavioural assessment instruments, the validation of the SCS is a complex task as self-control behaviours are mostly covert and must be inferred from the behaviour of persons under specific circumstances or from self-reports (Rosenbaum, 1980, p. 115). Convergent and discriminant validity of the SCS were assessed by comparing scores obtained on the SCS to scores obtained on a number of existing scales such as the Rotter's I-E scale, the Irrational Beliefs Test developed by Jones (1968) and the G factor (self-control) of Cattell's 16 Personality Factor Test (Rosenbaum, 1988, p. 493). Further to this, criterion related validity was established by comparing test scores with other experiments namely, various kinds of anti-depressant cognition and behaviours. Also, high correlations were obtained between SCS scores and assessment of specific self-efficacy expectations in situations that required self-control behaviour

(Leon & Rosenthal, 1984; Rosenbaum & Ben-Ari, 1985). From this it was concluded that the SCS is a valid instrument for assessing learned resourcefulness (Rosenbaum & Palmon, 1984, p. 245).

#### *5.2.6.9 Justification*

Rosenbaum (1980, p. 119) reports that the SCS may be useful to the researcher who wishes to assess individual differences in the tendency to employ self-control methods, as these methods refer to certain coping skills. For this reason, the use of the SCS supports the empirical aim which is to ascertain whether salutogenic constructs could differentiate between copers and noncopers in the sample of community service doctors. Furthermore, the SCS was found to be both a reliable and a valid instrument for assessing learned resourcefulness.

### **5.2.7 The Background Information Questionnaire**

The background information questionnaire is used to determine the nature of the population, with a view to constructing a profile of the individual (see appendix A).

#### *5.2.7.1 Development*

A background information questionnaire was developed by the researcher to investigate participant's ages, gender, marital status, number of children, religion, place of university qualification, other qualifications, employment location (rural or urban hospital) and hours worked per week. A self-report section was added in an attempt to investigate respondent's opinions about the role of social support and the adequacy of their academic training (see appendix A).

#### *5.2.7.2 Rationale*

It is accepted that the background questionnaire gives important information concerning the population of the study (Bless & Kathuria, 1993) and was, therefore, considered necessary for inclusion in this study.

#### 5.2.7.3 *Description*

The biographic part of the questionnaire consisted of 10 questions regarding the respondents: name, age, gender, marital status, number of children, educational qualifications, place of study, place of current employment and name of hospital where currently employed. This was followed by a short self-report section asking participants four questions, namely, the number of hours worked; their opinion on the value of social support, who/where support was received from and their opinion on the adequacy of their undergraduate training (see appendix A).

#### 5.2.7.4 *Administration*

The background information questionnaire is a self-completing instrument. For the self-report section, the respondents were asked to choose which option of a range of options he or she believed to be most descriptive of their current situation. There are no right or wrong answers and respondents were asked to answer spontaneously. There was no time restriction.

#### 5.2.7.5 *Scoring*

The background information questionnaire was analysed statistically. Descriptive statistics were made available and correctional analyses done on the respective variables.

#### 5.2.7.6 *Interpretation*

The interpretation for biographical data are obvious. The descriptive statistics for the biographic part of the questionnaire and self-report section proved useful in indicating the nature of the sample. Also, asking about the number of hours worked was an indication of workload.

#### 5.2.7.7 *Reliability*

There are only two main types of reliabilities, namely, internal consistency reliability and test-retest reliability (Hays, 1963). Firstly, in order to get a cronbach alpha value, a scale consisting of three or more items is required. There is thus no way to determine the internal consistency of (single) biographic variables. Secondly, the present research did not permit a test-retest reliability study. Although the reliability of biographical variables could not be determined in this research, it should be noted that the researcher can often accept such variables as reliable. With regards to a biographical variable such as “age of participant” for example, it is reasonable to assume that the respondent reports his or her age reliably. Such variables give important information concerning the population of the study which permits their inclusion in research.

#### 5.2.7.8 *Validity*

As with reliability, the validity of biographic variables is generally accepted. Following from the discussion of reliability above, only test-retest reliability of the data can affect the reliability which is a pre-condition for validity (Kerlinger, 1986). Since this is not a test-retest research, the validity of the biographical information questionnaire could not be determined, but is accepted on the basis of its obviously high face validity.

#### 5.2.7.9 *Justification*

Biographical questions are always asked in research, because practice has shown them to be important variables in defining a population.

This concluded step 2 of the empirical study.

### 5.3 DATA COLLECTION

Step 3 of the empirical study (see chap. 1, sec. 1.7.2), namely data collection, is addressed.

The quantitative assessment was conducted in October 1999 and lasted for a period of one month.

Permission was granted from supervisors to meet with the community service doctors. Altogether, 17 hospitals responded positively (10 from rural hospital and 7 from peri-urban hospitals).

The researcher met with and administered the seven quantitative instruments to 30 community service doctors. The remaining 11 participated in the study via the mail. Altogether, the sample consisted of 41 community service doctors ( $N = 41$ ).

The researcher opened the assessment session by introducing the aims and other details of the study – both verbally and in letter form. This was followed by an outline of the programme for the session (complete three questionnaires, lunch break, complete remaining three questionnaires).

The measuring instruments were administered in small group settings. The researcher ensured that all the participants were seated comfortably at a table with the relevant assessment material and stationery.

Participants were asked to complete the biographical information questionnaire. Respondents were informed that they could remain anonymous if they preferred and that all information would be treated confidentially. The participants were also informed that feedback on their results was available on request.

Instructions for completing each psychometric questionnaire (as discussed in sec. 5.2) were administered as recommended by the manual for the respective questionnaires. Some respondents needed further clarification. No time

restrictions were set. The entire session took approximately one and half hours to complete.

The answer sheets were hand scored and double-checked for accuracy.

This completes step 3 of the empirical research.

## **5.4 DATA PROCESSING**

Step 4 of the empirical study (see chap. 1, sec. 1.7.2), namely the statistical processing and discussion of the procedures and statistical techniques, will be addressed.

The statistical software package called Statistical Software Systems (SAS) was used for data processing.

The following data processing steps were performed.

### **5.4.1 Scoring and scaling of the measuring instruments**

Each of the scales was computed by determining the mean score of each item represented in that scale. To facilitate the comparison of the total scores of various scales, it was decided to transform scale scores for each respondent in the sample as follows:

- The stress and burnout scales were scaled so that a high score would indicate high stress and burnout. (However, for the subscale personal accomplishment on the Maslach Burnout Inventory, a high score on personal accomplishment indicated low burnout and a low score indicated higher burnout.)
- For the salutogenesis scales, all items were scaled such that a high score on the scale indicated “positive” salutogenesis”, so that all salutogenesis scales are in the same direction and could be interpreted in a similar way.

- The locus of control scale was scaled in such a way that a high score on the test indicated internal locus of control. Thus, positive correlations are expected with the salutogenesis variables and negative correlations are expected with the stress and burnout scales.
- The total score per instrument was calculated as the mean of the items comprising that scale.

To ease the comparison of the scale scores of the various measuring instruments, all scales were transformed to a scale ranging from 0 to 100 using the following formula:

$$\text{transformed score} = (\text{raw scale score} - \text{smallest score on scale}) / \text{range} \times 100$$

In this study, the hardiness of a subject could, for example, vary from 0 to 3 on any item. Since the total scale score for a subject was calculated as the mean of items, it implies that the total scale score can range from a minimum score of 0 to a maximum score of 3. The formula is thus

$$\text{transformed score} = (\text{raw scale score} - 0) / (3 - 0) \times 100$$

- The means and standard deviation scores of the normative samples (helping professions) for the stress, burnout and salutogenic instruments were converted to means and standard deviations on 0 (low) to 100 (high) scale to facilitate comparisons with the present sample. The formula for the transformation of mean and standard deviations are (see appendix B for details):

$$\text{transformed mean} = (\text{old mean} - \text{smallest value}) / (\text{possible range}) \times 100$$

$$\text{transformed standard deviation} = [\text{square root of } [(\text{variance} / (\text{square of range}))^*] \times 100 \text{ (here the variance is the square of the standard deviation)}$$

In cases where the standard deviation was not available, it was taken to be equal to that of the present sample for the sake of computing a *t*-test on the means; it was noted that this could affect the *p*-value slightly (motivation for using the *t*-test in this research is presented below in sec. 5.4.8).

#### **5.4.2 Reliability analysis of the measuring instruments**

For a large part, the present study assumes that the selected questionnaires are valid and reliable (see sec. 5.2). This research reports the internal consistency reliability, the cronbach alpha coefficient, for each of the measuring instrument scales used (Lemke & Wiersma, 1976). To assist with these reliability analyses, single factor solutions were extracted by means of factor analysis (Kerlinger, 1986). By considering the factor loadings of each individual item and the size and sign of the loading, the researcher can confirm the significance and intended direction of the scale of an item. This proves useful when decisions have to be taken as to which items should be omitted from the scale (if any at all), and which items should be re-scaled, which items should not be re-scaled.

In the present study, items with factor loadings smaller in absolute size than 0,10 were omitted. This cut off value of 0,10 is rather conservative as often a more stringent cut of - of 0,25 or even 0,30 is used (Field, 2000, p. 440). Since most measuring instruments are established instruments accompanied by a small size, it would be prudent to attempt to retain as many of the items in a particular scale as possible.

Single factor solutions were obtained for the dimensions and subdimensions of the various psychometric questionnaires administered.

#### **5.4.3 Level of stress**

Descriptive statistics (means, standard deviations, frequencies and percentages) were calculated for each subscale of the Stress Diagnostic Survey (SDS), namely role conflict, role ambiguity, work overload (quantitative and qualitative), career development and responsibility for people. Each of these five individual level

stressor categories was classified into low, moderate or high stress according to the sum of their item scores.

A chi-square test was performed to examine the significance of the association between levels of stress and the SDS scales. The chi-square statistic is useful for comparing frequencies and identifying the extent to which differences in frequencies are reliable and not the result of chance events or sampling (Clifford, 1981 in Arumugam, 1992).

The means and standard deviations were compared to the mean and standard deviations of the SDS comparative sample. Based on this, *t*-scores were calculated to assess whether the sample was experiencing significant levels of stress or not (motivation for using the *t*-test in this research is presented in sec. 5.4.8). According to Bless and Kathuria (1993), the aim of the *t*-test analysis is to compare the means of the two groups, that is, to analyse their difference on the assumption that the two groups belong to the same population or to two populations with the same mean.

#### **5.4.4 Level of burnout**

Descriptive statistics (means, standard deviations, frequencies and percentages) were calculated for each Maslach Burnout Inventory (MBI) subscale, namely emotional exhaustion intensity and frequency, depersonalisation intensity and frequency, personal accomplishment frequency and intensity.

Percentages were also compared to the normative group for each subscale. In this way, every MBI subscale was categorised into high, moderate or low burnout groups. This is in keeping with Maslach and Jackson's (1981) conceptualisation of burnout as a continuous variable, ranging from low to moderate to high levels of experienced feeling, not as a dichotomous variable either absent or present.

The means and standard deviations were compared to the mean and standard deviations of the MBI normative sample. Based on this, *t*-scores were calculated

to assess whether the sample was experiencing significant levels of burnout or not (motivation for using the *t*-test in this research is presented below in sec. 5.4.8).

#### **5.4.5 Level of salutogenic functioning**

Descriptive statistics (means and standard deviations) were calculated for each of the four salutogenic measuring instruments (sense of coherence, hardiness, locus of control and learned resourcefulness) and their respective subscales.

As previously indicated, a total score was calculated for each subject on each of the salutogenic scales by calculating the subjects' mean score over the items of the scale and then transforming the scale to a range from 1 to 100.

The mean and standard deviations of the normative samples (mostly helping professions) for each salutogenic instrument were converted to means and standard deviations on a 0 (low) to 100 (high) scale to facilitate comparisons with the present sample (see section 5.4.1 for the formula for these conversions).

#### **5.4.6 Correlations between stress, burnout and salutogenic variables**

In the case when two variables are both continuous (eg locus of control scale and the salutogenic subscales), the appropriate statistic to compute is the Pearson's Product Moment Correlation statistic (Hays, 1963). This statistic varies from  $-1$  (a perfect negative linear relationship) through zero (no linear relationship) to  $+1$  (a perfect positive linear relationship). Correlations were interpreted in the following manner:

- If the correlation is negative between two scales, this indicates that the higher the scores on the one scale, the lower the scores tend to be on the other scale.
- If the correlation is positive, then the higher the scores on one variable, the higher the scores on the other variable tend to be.

The derivation of composite scores is discussed next.

#### **5.4.7 Calculating composite scores**

In the present study, a number of scales were used for each major construct, namely stress, burnout and salutogenesis. A factor analysis was performed on all the scales measuring salutogenic functioning and a single factor solution forced and the factor loadings studied in order to judge which of the individual scales may be used to derive a composite salutogenic variable. The same approach was followed with regard to the various scales measuring stress and the scales measuring burnout.

To determine whether the salutogenic functioning scores could differentiate between copers and noncopers, the composite score Burnout created above was of special importance for the purpose of this research. In this regard, the following assumption is made: higher burnout scores are equated with noncoping, whilst lower burnout scores are equated with coping. This assumption has been justified in the literature review integration (end of chap. 4).

Based on the composite scores, the respondents' overall performance was prioritised in terms of stress, burnout and salutogenic functioning. Descriptive statistics were also calculated for the three composite variables. With respect to the composite variable Burnout, using the mean and sample size as criteria, cut off points were established and the group was divided into high and moderate burnout groups. The group that scored in the midrange was ignored, because the aim was to compare two groups which were as extremely different as possible on burnout and could thus be considered a group of copers and a group of noncopers.

The motivation for using factor analysis in this research is presented next.

**Note: Motivation for using factor analysis considering sample size**

In the present study, factor analysis was used only in a limited way since - at most - a single factor was extracted that did not require the rotation of factor solutions. The use of factor analysis was also restricted in the sense that it was never used in a confirmatory-hypothesis testing sense, which would have required assumptions about the form of the multivariate distribution of item/test scores.

The factor loadings of the single factor solutions were mostly simply inspected to assist the researcher in ensuring that items that constituted a particular scale, did score in the correct direction. Items that load negatively must, therefore, be reverse scored so that they measure the underlying construct in the same direction as the remaining items, or the items loading positively must be reverse scored to measure in the same direction as the items loading negatively. Apart from using factor analysis to assist in the item analysis of the various scales, factor analysis was also used to derive single factor solutions for the subscales of stress, burnout and the various subscales of salutogenesis.

Arrindell and Van der Ende (1985, p. 175) suggested, “for the purpose of establishing a given number of true factors, sample size should be related to the number of factors drawn”. These authors show that in order to obtain stable factor solutions, the sample size should be “approximately 20 times the number of factors” (Arrindell & Van der Ende, 1985, p. 175). For example, in the case of a single factor, the sample size requirement would be  $20 \times 1 = 20$ . As the present research used a sample size of 41, this was considered sufficient for the limited purposes for which factor analysis was used in the present study.

**5.4.8 Establishing statistical significance**

Finally, in order to determine whether significant differences existed between the salutogenic functioning scores and the respondents who scored high and moderate on burnout, the *t*-ratio was determined with a *t*-test (motivation for using the *t*-test is presented below). According to Kerlinger (1986, p. 207), the *t*-test is the central statistic behind the analysis of variance, and the main aim of this

procedure is to determine whether the means differ significantly. The motivation for using the  $t$ -test is presented next.

**Note: Motivation for using the  $t$ -test (the parametric  $t$ -test versus the non-parametric Mann-Whitney test).**

The  $t$ -test for independent measures tests whether two groups differ with respect to their mean scores on some interval or ratio scale. The Mann-Whitney test, on the other hand, is a non-parametric test that tests whether two independent groups have different distributions. If, in the latter instance, significant differences are found, one still does not know how to interpret such differences. Boneau (1960) has shown that in the case of the  $t$ -test for independent samples, "... sample sizes of 15 are generally sufficient to undo most of the damage inflicted by violations of assumptions" (Boneau, 1960, p. 60).

Since the groups being compared with respect to means in the present study were not smaller than 15, it was decided to use the  $t$ -test to compare independent groups with regard to their mean scores. Furthermore, in most cases where the present sample was being compared against a norm group (data from research literature or manuals), the full 41 cases were used.

In some cases, the researcher could only find mean scores for norm groups and no standard deviations. In such cases, the standard deviation of the present sample was taken to be the same for the norm group, so that a  $t$ -test could be computed (it is noted that this may affect the  $p$ -values slightly).

#### **5.4.9 Level of statistical significance**

Conventionally, the levels 0,05 and 0,01 are used by most researchers as levels of significance for statistical tests performed (Hays, 1963; Winer, 1971). In choosing a level of significance for the present research, the following viewpoints were taken into account:

- In the human sciences, the concern with missing a significant result or making a type-II error is as crucial as the concern about falsely concluding a significant result. Hays (1963) and Winer (1971) point out that when both types of errors (type I and type II) are equally important, significance levels such as 0,20 (and possibly even 0.30) are more appropriate than the conventionally used 0,05 and 0,01 levels.
- As the total number of statistical tests to be performed on the same sample increases, the probability of a type I error also increases. One approach to counter this accumulating effect is to set the level of significance smaller for the individual statistical test so as to compensate for the overall type I error effect. If the overall research significance level is 0,30, for example, then the significance level for the individual test might be 0,05 or 0,01. There is no easy way to decide exactly what this level of significance should be.
- The sample size in the present study is on the small side ( $n= 41$ ) so that statistical tests lack power, that is, the ability to detect significant results (Hays, 1963). In other words, it is difficult to detect significant results.

In view of all these considerations, a significance level of 0,05 for any one particular statistical test in the present research was chosen.

Finally, it was ensured that all the research questions had been answered.

This concludes step 4 of the empirical study

## **5.5 HYPOTHESES**

The central research hypothesis underlying this research was outlined in chapter one.

With reference to step 5 (chap. 1, sec. 1.7.2) of the empirical study, the integration of the main variables (see end chap. 4) indicated that certain relationships existed

between stress, burnout and salutogenic functioning. The purpose of the central thesis is thus to confirm this relationship, empirically, and to indicate to what extent this relationship actually exists. The central thesis, serving as guideline, for this research is thus as follows:

“There exists a relationship among the community service doctors’ salutogenic functioning and their stress and burnout levels in KwaZulu-Natal hospitals.”

The following hypotheses were formulated from this central research hypothesis:

### **Hypothesis 1**

The community service doctors are experiencing higher levels of stress than other helping professions.

### **Hypothesis 2**

The community service doctors are experiencing higher levels of burnout than other helping professions

### **Hypothesis 3**

The community service doctors have lower levels of salutogenic functioning than other helping professions.

### **Hypothesis 4**

The lower the level of salutogenic functioning in community service doctors, the higher the stress these doctors will experience.

### **Hypothesis 5**

The lower the level of salutogenic functioning in community service doctors, the higher the burnout these doctors will experience.

## **Hypothesis 6**

Salutogenic functioning scores can differentiate between copers and noncopers in community service doctors, whereby copers will show higher levels of salutogenesis than noncopers.

This concludes step 5 of the empirical research.

## **5.6 CHAPTER SUMMARY**

In this chapter, steps 1 to 5 of phase 2 of the research methodology (see chap. 1, sec. 1.7.2) were discussed. The chapter opened with a description of the sample (step 1). This was followed by a discussion of the measuring battery (step 2). Here the quantitative instruments were explained in terms of its development, rationale, description, administration, analysis, interpretation, reliability, validity and justification of its inclusion. Setting the scene for the results in chapter 6, the data collection and statistical processing were then discussed in terms of the empirical objectives (step 3 and step 4). Finally, the chapter concluded with the formulation of the hypotheses of the research (step 5).

### **▪ REMARK**

With reference to the research methodology in chapter 1, section 1.7:2, phase 2, steps 1-4, the objective of the empirical study have been established.

The next chapter will look at reporting, interpreting and integrating the results of this research.