TRANSLATION, PSYCHOMETRIC EVALUATION, AND PRELIMINARY VALIDATION OF A ZULU VERSION OF THE SCL-90-R

SEAN FRANCIS SHANAHAN

Submitted in partial fulfilment of the requirements
for the degree of
Master of Social Science (Clinical Psychology)
in the
Department of Psychology
University of Natal
Pietermaritzburg

January 1998

ABSTRACT

This study is seen as a step towards addressing the acute shortage of psychometric instruments that can be validly and reliably used in the assessment of Zulu-speaking individuals. The Symptom Checklist-90-Revised (SCL-90-R), a 90-item multidimensional self-report symptom inventory, was translated into Zulu. A multistage translation procedure, involving back-translation, decentering, and the committee approach was employed. The translated instrument was pretested on a group of Zulu farm workers (N = 12) and revisions made in order to improve its comprehensibility and acceptability to Zulu respondents. The psychometric equivalence of the Zulu and English versions of the SCL-90-R was investigated with bilingual Zulu students (N = 61). Scale-level factor analysis yielded similar factor structures for both versions, and suggests that the Zulu SCL-90-R may be best utilised as a global measure of psychological distress. Scores at the scale- and item-level were reasonably comparable for the two language versions, although retest effects and apparent bilingual response sets suggest that the bilingual technique may not be a valid means of assessing translation equivalence. Acceptable test-retest reliability and internal consistency measures were obtained, indicating that the translation into Zulu did not adversely affect the reliability of the SCL-90-R. The concurrent validity of the Zulu SCL-90-R was investigated with samples of male psychiatric inpatients (N = 23) and nonpatients (N = 26). The global severity index of the SCL-90-R demonstrated moderate diagnostic efficiency, with a sensitivity of 70% and a

specificity of 77%. These results suggest that the Zulu SCL-90-R may be validly utilised for the purpose of screening for mental illness.

DECLARATION

This thesis was undertaken at the Department of Psychology, University of Natal, Pietermaritzburg, and, unless specifically indicated to the contrary in the text, is a product of the author's own work. This thesis has not been submitted to any other university.

Sean Francis Shanahan
January 1998

ACKNOWLEDGEMENTS

This research was made possible by the generous assistance of numerous people. I am particularly grateful to the following:

Dr Stuart Anderson, my supervisor, for suggesting this area of study to me, and for his assistance and advice over the duration of this project.

Mr Nhlanhla Mkhize, who co-supervised this research, for his critical guidance during the translation phase.

Dr Bruce Faulds, for his generous assistance with the SPSS data analysis, and his supervisory guidance during Dr Anderson's period of absence.

The translation team, Jabulani Thwala, Sibu Ntshangase, Elliot Makhathini, Baba Hlengwa, Mary Gordon, and Nelson Ntshangase.

Dr Walker of Midlands Hospital Complex, Baba Hlengwa of the Department of Zulu, Mr Joe Meyer of the Ukulinga Research Farm, and Mr Warren Forsyth of the University of Natal Estates Division, for granting permission to conduct research at those institutions.

The patients, students, and workers who participated as subjects in this study.

Dr Paul Shanahan, for assistance with the graphs.

The Centre for Science Development, for financial assistance. The views expressed in this work are those of the author and are not necessarily shared by the Centre for Science Development.

Special thanks is due to my parents for their patience and assistance; and to my partner, Margaret, for her loving support and encouragement.

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In the interests of developing appropriate mental health services to meet the needs of all South Africans, it is essential that the limited available resources be optimally employed. It is widely recognized that existing mental health services are grossly inadequate and that the realistic restructuring of such services needs to be based on reliable epidemiological data (Freeman, 1992; Miller & Swartz, 1992; Subedar, 1993). Consequently, the lack of suitable psychometric instruments, which may be validly used for gathering data on mental health issues in a multi-cultural and multi-linguistic society, is a problem which must be addressed with some urgency.

A related issue concerns the need for instruments which may be used for the screening of mental illness and psychological distress at primary health care settings in South Africa. It has been estimated that depressive and anxiety disorders alone account for 20-30% of all primary health care visits worldwide (Desjarlais, Eisenberg, Good, & Kleinman, 1995). The vast majority of these cases are not correctly diagnosed and are consequently denied effective treatment.

The current research is seen as a step towards addressing the acute shortage of psychological instruments which may be validly used in the assessment of Zulu-speaking individuals, Zulu being the mother-tongue of the majority of the population of the province of KwaZulu-Natal. The goals of the project are as follows:

a) to produce a rigorously translated Zulu version of the SCL-90-R;

b) to assess the psychometric equivalence of the Zulu translation and the original SCL-90-R; and

c) to conduct a preliminary validation study to ascertain whether the Zulu version SCL-90-R is capable of distinguishing between psychiatric patients and nonpatients.

The SCL-90-R (Derogatis, 1983) is a multidimensional, 90-item, self-report symptom inventory designed to measure symptomatic psychological distress in a range of individuals, including psychiatric patients, medical patients, and normal nonpatients. The instrument has been subjected to extensive study, and has been translated into at least 26 languages (Derogatis & Lazarus, 1994). However, according to the publishers (NCS Assessments, personal communication, 1996), no Zulu version of the SCL-90-R currently exists.

The SCL-90-R satisfies all the criteria suggested by John (1996) for the selection of instruments to be adapted for use in another culture:

a) the SCL-90-R has been extensively researched and its psychometric properties adequately demonstrated in the country of origin, the United States;

b) the instrument is embedded in a substantial body of literature;

c) the literature points to the successful application of the instrument in other cultures; and

d) the instrument is cost-effective, being easy to administer and score, and not requiring expensive test materials.

The use of instruments that have been conceived, standardised, and validated in Western cultures in the assessment of individuals from other cultural backgrounds is a contentious issue (Ebigbo, 1982; Gordon & Kikuchi, 1966; Malgady, Rogler, & Constantino, 1987). In particular, there is a danger of employing an "imposed etic" (Triandis, 1980). This is when culturally specific (emic) aspects of the source culture are applied to the target culture as if they were universally

applicable (etic), while important emic aspects of the target culture may be neglected. Means of attempting to avoid such imposition include close collaboration, during all phases of the test adaptation process, with colleagues who are members of the target culture; careful pretesting of the instrument with subjects who are representative of the target culture; modification of items which are found to be culturally inappropriate; and, if necessary, the addition of items to measure important emic aspects of the target culture. Finally, the reliability and validity of the test, as applied to the target culture, must be empirically demonstrated.

The advantages of translating an established instrument such as the SCL-90-R include the low costs involved, relative to attempting to construct a new, indigenous instrument, and the availability of an accumulated body of research supporting the conceptual and psychometric properties of the original instrument. Data obtained with the translated instrument may be compared and interpreted in the light of this research. Another pragmatic reason, suggested by Cheung (1985), is that an established instrument may be more readily accepted and utilised by local mental health practitioners and researchers. Further, if the SCL-90-R can be successfully adapted for use with Zulu-speaking individuals, the likelihood exists that the instrument may be adapted for use with other language groups in South Africa. Restructuring of health care services on a national level may be best served if epidemiological data for the country's various language groups are gathered with comparable instruments.

The current study recognizes the importance and complexity of translation as applied to the area of cross-cultural psychological assessment. Several methods of translation are combined in a multiphase procedure which views test adaptation as a process involving a series of evaluations and refinements. The following chapter considers the problem of translation as applied to social scientific research, and reviews the available

translation methods. Chapter 3 describes and critically discusses the SCL-90-R, and reviews the research relating to the instrument's reliability, validity, and factorial structure. Chapter 4 reports the process of translating the SCL-90-R into Zulu, the problems encountered, and the attempted solutions to those problems, while Chapter 5 reports on the pretest study and the ensuing revisions to the Zulu SCL-90-R. Chapter 6 presents and discusses the results of a psychometric evaluation of the Zulu SCL-90-R, based on the responses of bilingual Zulu students. Finally, Chapter 7 reports on a preliminary validation study, considering whether the Zulu SCL-90-R can be validly utilised to distinguish between psychiatric patients and nonpatients.

CHAPTER 2

TRANSLATION CONCEPTS AND METHODS

This section considers the problem of translation as applied to social scientific research, and reviews the available translation methods. Translation, as defined by Brislin (1976), is the transfer of thoughts and ideas from one language to another. In the area of cross-cultural assessment, this transfer of thoughts and ideas is arguably both the most important and most complex task facing the researcher. The importance of the task can be seen in the fact that the quality of the translated instrument must impact on each ensuing phase of a study and will ultimately determine the quality of the data collected. In particular, in the absence of a systematic and rigorously applied translation procedure, the alternative hypothesis of data contamination due to translation error cannot be excluded. The complexity of the task is reflected in the hyperbolical assertion by Richards (as cited in Brislin, 1980, p. 426) that translation is "probably the most complex type of event yet produced in the evolution of the cosmos," and Brislin's (1976) characterisation of the translation process as "an example of an adverse, difficult state of affairs" (p. 8).

While the difficulties associated with translation have led Triandis (1972, 1976) to propose methods aimed at minimising the use of translation in cross-cultural research, many other researchers (Bravo, Canino, Rubio-Stipec & Woodbury-Farina, 1991; Brislin, 1970, 1976, 1980, 1986; Brislin, Lonner & Thorndike, 1973; Butcher, 1982; Butcher & Clark, 1979; Butcher & Pancheri, 1976; Geisinger, 1994; Prince & Mombour, 1967; Retief, 1988; Sechrest, Fay & Zaidi, 1972; Sperber, Devellis & Boehlecke, 1994; Werner & Campbell, 1970) have provided guidelines to assist in the translation of assessment instruments. According to Brislin (1976) the "most careful and formal checks on translation quality

have been designed (and used) for empirical research in the social and behavioral sciences" (pp. 19-20).

Given the availability of a fairly extensive literature on translation methods, and considering the central importance and complexity of translation in cross-cultural research, the general lack of attention to translation issues in cross-cultural research reports is remarkable. The need for researchers to be more rigorous in the application of available translation methods, and for the process of such translation efforts to be reported, has been repeatedly emphasised (Brislin, 1970; Brislin et al., 1973; Butcher & Clark, 1979; Prince & Mombour, 1967; Rogler, 1989; Sperber et al., 1994; Werner & Campbell, 1970). Reporting on a review of 80 articles in major cross-cultural psychological journals, Brislin (1970) noted that approximately 75% gave little or no information concerning the translation methods used. More recently, Rogler (1989) states that despite improvements in available translation procedures "it appears doubtful that there has been a substantial increase in general efforts towards the goal of culturally adapting psychiatric research" (p. 299). Reviewing the South African research literature, Drennan (1992) and John (1996) similarly comment on the lack of attention given to translation issues, and the general failure of South African researchers to adequately utilise the applicable techniques.

2.1 TYPES OF TRANSLATION PROBLEMS

Drawing on the work of Brislin (1970) and Werner and Campbell (1970), Sechrest et al. (1972) identify four broad types of translation problems facing the cross-cultural researcher. They report that the attention given to translation in these areas has been selective and insufficient. The first type of translation problem relates to the *orientation of research subjects* to the research being conducted. Research work typically requires that some sort of rationale be given to

subjects regarding the tasks that they are expected to perform. Research ethics necessitate that respondents understand the purpose of the research and how the results will be used (Pareek & Rao, 1980). It is consequently necessary that attention be given to attempting to ensure the equivalence of such explanations across cultures. Lack of due attention to this variable may result in confused subjects and biased responses (Retief, 1988).

The second type of problem involves the translation of specific *instructions* relating to the particular tasks or measures being used. Irvine (1986) cautions against assuming that subjects understand testing conventions, and emphasises the importance of clear instructions. Werner and Campbell (1970) point out that researchers should resist the tendency to assume equivalence of instructions, particularly when the instructions are kept brief. Lack of context and lack of redundancy in brief instructions make the attainment of satisfactory translation of meaning more difficult than may be the case with longer instructions.

The third type of translation problem is the obvious requirement of phrasing the actual *questions* or other verbal stimuli in ways that are comparable in the two (or more) languages. Examples include interview questions, items on symptom checklists, statements on personality inventories or personality questionnaires, and verbal stimuli on projective techniques such as incomplete sentences. Sechrest et al. (1972) make the comment that while this problem has received more attention from researchers than the other three types of translation problems, many investigators have paid insufficient attention to the complexities involved. These complexities include the achievement of various types of equivalence, such as idiomatic equivalence and experiential equivalence, which are discussed below.

The final type of translation problem involves the translation of subjects' *responses*. In the case of responses arising from open-ended questions, interviews, projective techniques, and so forth, the actual responses given by the subjects may need to be fully translated prior to analysis. Alternatively, the responses may be coded, categorised, or scored in the language in which they are given (Sechrest et al., 1972). In the case of questionnaires with limited response alternatives, such as true-false or likert scale categories, the translation is made prior to the response, and due attention should be paid to the appropriateness of the categories in the target culture. This is clearly an important aspect of test translation; culturally inappropriate or poorly translated response categories will affect the responses to each and every item. For example, likert scale categories (such as "not at all; a little bit; moderately; quite a bit; extremely") may be difficult to translate into a target language in which similar graduated distinctions do not exist or are not commonly used. If the instrument is to be used for cross-cultural comparison, it is necessary to ensure that respondents have comparable understandings of the scale values or distances in each culture (Pareek & Rao, 1980).

2.2 GOALS OF TRANSLATION

Brislin (1980) regards a consideration of the different purposes of translation as a useful start to any discussion of translation. He refers to, and expands on, Casagrande's (cited in Brislin, 1980) four types, or "ends," of translation. The first type, *pragmatic translation*, is concerned purely with the accurate communication of the factual information that was meant to be conveyed in the source language form. An example of this type is in the translation of technical documents, such as machine repair manuals, where subtleties of language (e.g., aesthetic form) are of no concern.

In contrast to this first type, *aesthetic-poetic translation* focuses on conveying the feelings, emotional meaning, and aesthetic form of the source language version, as well as any factual information contained in the message. The translation of creative literature is an example of this type.

The third type is *ethnographic translation*, the goal of which is to interpret and explain the cultural contexts of the source and target language versions. As stated by Werner and Campbell (1970), "translation also means providing an understanding of cultural empathy, where content, context and style may have to be adjusted to the receptor culture" (p. 399). This is potentially a highly complex task, requiring bilingual competence as well as an intimate understanding of social arrangements and sensitivity to the various connotations of words used in the contexts of the cultures being considered. Towards this end, Retief (1988) emphasises the importance of collaborators who are familiar with both the source and target cultures as well as their languages.

The fourth type, *linguistic translation*, is concerned with achieving correspondence of semantic meaning as well as linguistic structure and grammatical form. This requires an analysis of the functional relations of the constituent morphemes of both the source and target languages.

Brislin (1980) notes that any single translation task can rarely be categorised into only one of the four types. He states, however, that awareness of these goals should be useful in assisting the translator in deciding on specific goals for a translation, and in making compromises and choices between competing objectives subsumed under the four types.

Retief (1988) and John (1996) regard the pragmatic and ethnographic types as being particularly pertinent to the cross-cultural translation of assessment instruments. Thus, the information contained in the source language is communicated in the target language, while "attention [is] given to ensuring that

the words chosen for the translated version are harmonious with the culture that speaks that language" (John, 1996, p. 60). This is at variance with Brislin et al. (1973), who regard psychological translation as entailing pragmatic translation, since the empirical demonstration of equivalence is implicit in the methodology of cross-cultural research, the other three types (including ethnographic translation) being less central.

Other researchers (Bontempo, 1993; Drennan, 1992; Hulin, 1987) have questioned the applicability of Casagrande's types to psychological research. Hulin (1987) contends that while mood and feeling are involved as is the case with aesthetic-poetic translation, the purpose of psychological scales is not to *evoke* moods and feelings, but rather to facilitate the reporting of the respondent's experience. With regard to linguistic translation, he argues that while equivalence in the structure of translated materials is required, the goal is to achieve equivalent stimuli rather than equivalent structure as an end in itself. In rejecting Casagrande's typology as applied to psychological scales, Hulin (1987) and Bontempo (1993) emphasise the need for the equivalence of item meanings and item responses in different language versions to be psychometrically demonstrated, a position broadly similar to that of Brislin et al. (1973).

Drennan (1992) criticises Casagrande's types as constituting "a confusing array of translation goals" (p. 3) which are not appropriate for the purposes of psychological research. An analysis of different types of equivalence is proposed as an alternative approach to examining the purposes of translation.

2.3 THE PROBLEM OF EQUIVALENCE

According to Sechrest et al. (1972), ensuring the equivalence of the translated material to the original language version is the central purpose of translation. They provide a useful discussion of several different aspects of equivalence

which should be considered when attempting the cross-cultural translation of research instruments.

2.3.1 Vocabulary Equivalence

This is probably the most obvious type, involving equivalence of the actual words used. If the words of test items do not have equivalent meaning to respondents in different cultures, the interpretation of responses is tenuous and cross-cultural comparisons cannot be made. Sechrest et al. (1972) make the point that, while a good dictionary is a valuable resource, translation based purely on dictionary consultation is not sufficient to achieve vocabulary equivalence. The language contained in dictionaries is often not the language used by the people for whom the translated material is intended, and may consequently be misunderstood by test respondents. Similarly, translators may differ markedly from the target population in terms of level of education and may consequently use a rather stilted version of the target language. With reference to the South African context, Retief (1988) makes the point that the use of scholarly rather than vernacular language in questionnaires renders relatively simple questions more difficult to understand. It is therefore essential that translators have a thorough knowledge of the language as commonly used by the intended test respondents.

There may be terms in one language for which equivalent terms do not exist in another language. A possible solution, recommended by Werner and Campbell (1970), is the use of a short description in the target language in an attempt to convey a concept which may be expressed by a single word in the source language. Differences in the length of materials should, however, be kept within reasonable limits (Sechrest et al., 1972).

Brislin (1970) and Sechrest et al. (1972) suggest that, as a result of cultural differences and divergent linguistic traditions, vocabulary equivalence may be more difficult to achieve between certain language pairs than others. In the South African context, a number of researchers (Buntting & Wessels, 1991; Gillis, Elk, Ben-Arie & Teggin, 1982; Le Roux, 1973) have commented on the difficulties encountered in translating English-language psychological materials into Black African languages.

2.3.2 Idiomatic Equivalence

Idiomatic expressions complicate the translation task and typically do not withstand direct or literal translation. It is for this reason that Werner and Campbell (1970) and Brislin (1980, 1986, Brislin et al., 1973) recommend that metaphors and colloquialisms be avoided for psychological research purposes. However, metaphor pervades all forms of discourse (Dunnigan, McNall & Mortimer, 1993) and most widely used psychological inventories (including the SCL-90-R) contain examples of idiomatic language. It is also noted that the Zulu language is particularly rich in idiom (Edwards, Cheetham, Majozi, & Lasich, 1985). Sechrest et al. (1972) argue that complete avoidance of idiomatic terms tends to produce laboured and pedantic phraseology which is unsuitable for research purposes with the general population of any culture. They suggest that if idioms are retained in a translation, the best that can be done is to attempt, with the assistance of bilingual consultants, to find idioms in the target culture which are of equivalent meaning and roughly equivalent familiarity (or frequency of usage) across the two languages. Butcher (1982) suggests the addition of an explanation, in parentheses, to add clarity to items containing figurative expressions.

While asserting that culturally different metaphorical representations should not be interpreted as necessarily indicating qualitative differences in the experiences of

different groups (cf., Whorf, 1956), Dunnigan et al. (1993) question whether research relating to highly metaphorised cultural domains (such as emotional states) can produce comparable data for culturally divergent groups. They believe that this question can only be answered via semantic analysis on a case-by-case basis, and point out the inadequacies of direct translation and back translation as applied to metaphorical language. Excellent examples of item analyses are provided by Butcher and Pancheri (1976, particularly chapter 3) and Clark (1985). In the African context, Buntting (1988), Buntting and Wessels (1991), Ebigbo (1982), Gillis et al. (1982), and Orley and Wing (1979) also provide useful item-level discussions. Buntting and Wessels (1991) refer to the greater difficulty encountered in the translation of items involving emotional states, as opposed to psychotic states, from English into Zulu.

In cases where, as a result of contrasting cultural traditions of metaphorical expressions, semantically equivalent assessment instruments are found not to be achievable by means of translation, Dunnigan et al. (1993) suggest that separate sets of culturally appropriate questions may have to be developed for each group. This position is consistent with that of Triandis (1972, 1976).

2.3.3 Grammatical-Syntactical Equivalence

Languages differ in their grammars and syntaxes, and these differences present problems for translators. For example, parts of speech, such as adjectives and adverbs, may differ in their function or may even be missing in certain languages. Guidelines for dealing with these problems are difficult to specify since different difficulties will arise according to the particular languages being used.

The practical relevance of this type of equivalence has been questioned in the literature. Due to differences in the function

and even existence of parts of speech across different languages, Drennan (1992) describes grammatical-syntactical equivalence as an unrealistic goal which may not have much utility in practice. Similarly, Retief (1988) asserts that the effective transmission of meaning should take precedence over excessive concern with grammatical and syntactical considerations which could lead to long-winded and confusing language.

This is not to say that grammatical considerations are not important. Butcher and Clark (1979) caution that ungrammatical translation may result in items being misunderstood, subjects losing confidence in the instrument, and consequent loss of proper test-taking attitude.

2.3.4 Experiential Equivalence

Translations from one culture to another should make use of terms which refer to actual things and experiences of approximately equal familiarity to both cultures. Werner and Campbell (1970) refer to this as "cultural translation" as opposed to those aspects of translation which relate to purely linguistic considerations. This type of equivalence is difficult to achieve to the extent that cultures differ in the nature of their physical environments, social arrangements, and so forth, or to the extent that things and experiences familiar to one culture are unfamiliar to another. While in certain cases it may be necessary to eliminate an item because a counterpart does not exist or would be of uncertain equivalence in another culture, the task of the translator is to ascertain what the item was intended to reflect in terms of a trait or response disposition in the source culture and then to find an equivalent which would be meaningful to members of the target culture (Sechrest et al., 1972). For example, Gillis et al. (1982) question the relevance of inquiring into weight loss in the assessment of African people who tend not to be weight conscious and may not have access to

scales. They suggest that a question concerning looseness of clothing may be an appropriate alternative in this case.

2.3.5 Conceptual Equivalence

A further problem of equivalence involves the question of whether the concepts which the translated materials are intended to measure are equivalent in the cultures being studied. Sechrest et al. (1972) describe two potential problems in this regard. Firstly, it may occur that a particular word or item may appear to translate adequately in terms of vocabulary equivalence, and yet fail to achieve equivalence of the underlying concepts implied by the word. For example, Butcher and Clark (1979) discuss the concept of shyness as used in the MMPI item "I wish I were not so shy." While such an item may be relatively easy to translate on a word-for-word basis, in many non-Western cultures the word for "shy" has a more positive connotation. Respondents from such cultures may consequently interpret the item "I wish I were not so shy" rather differently to respondents from the source culture. As expressed by Prince and Mombour (1967), "it is the problem of connotation, of secondary meanings, that presents the gravest problem. Words are focal points of complex networks of meaning. They receive shades of significance from the most diverse and unexpected sources: clang associations, rhymes, associations by historical accident, linkages with social class, literary and journalistic accretions and a host of other sources" (p. 230).

Secondly, a concept familiar to one culture may be completely foreign to another, as indicated by the absence of an equivalent term in the target language, or may appear in varied or fragmented forms in the target language. Back-translation together with consultation with bilingual experts may be a means of identifying and seeking solutions to this type of conceptual nonequivalence. Potential solutions include revising the original item (see discussion below on decentering) or omitting

the item. Retief (1988) suggests that in some cases where a suitable equivalent does not exist in the target language, a word may be retained in the original language if it is commonly used by the target population.

Sechrest et al. (1972) and Poortinga (1989) make the important point that cross-cultural conceptual nonequivalence should not be seen purely as an inconvenient complication in the translation process, but may offer insights into important cultural differences. These insights should be incorporated into the ongoing research process.

Flaherty, Gaviria, Pathak, Mitchell, Wintrob, Richman, and Birz (1988) offer an alternative approach to the problem of equivalence. They propose five major dimensions of cross-cultural equivalence which constitute a stepwise validation process for instrument development. Three of these dimensions (content, semantic, and conceptual equivalence) overlap with those of Sechrest et al. (1972). The remaining two dimensions (technical and criterion equivalence) go beyond Sechrest et al. (1972) and are reviewed here.

2.3.6 Technical Equivalence

This pertains to whether the method of data collection affects the results differently in the two cultures. For example, in groups characterised by limited schooling and high illiteracy rates, the use of paper and pencil data collection methods is problematic. Similarly, as discussed by Gillis et al. (1982), an interrogative style of interviewing is foreign to the expectations and practices of many traditional African groups, and may be experienced as more offensive by certain subject groups than by others. Draguns (1982) regards cultural variation in the willingness to share subjective experiences, particularly in the standardised and impersonal manner required by psychological scales, as probably the most basic problem with the

application of such instruments. Irvine and Carroll (1980) caution against the assumption that subjects understand testing conventions, and provide guidelines for cross-cultural test administration. These include giving instructions verbally; the use of examples to demonstrate types of items and responses; allowing subjects supervised practice in order to ensure that they understand the test materials and method of responding; and the employment of trained test administrators of the same cultural group as the respondents.

2.3.7 Criterion Equivalence

The interpretation of the results of the measure should remain the same when compared with independent normative criteria for each culture. This refers to the concepts of predictive and concurrent validity. For example, an instrument designed to measure depression should be shown to distinguish depressed from nondepressed groups within each culture based on previously established criteria for depression. This type of equivalence recognises the importance of cultural values in the diagnostic process and seeks to avoid the "category fallacy" (Kleinman, 1987). The category fallacy is defined as "the reification of a nosological category developed for a particular cultural group that is then applied to members of another culture for whom it lacks coherence and its validity has not been established" (ibid., p. 452). As stated by Wakefield (1992, pp. 383-4), "[t]o be considered a disorder, the dysfunction [e.g., hallucinations] must ... cause significant harm to the person under present environmental circumstances and according to present cultural standards." In the African context, for example, the assessment of psychotic symptomatology should take account of, and not pathologise, culturally consonant experiences such as those associated with *ukuthwasa*, the calling to become a traditional healer.

A number of other researchers (Berry, 1980; Berry, Poortinga, Segall, & Dasen, 1992; Butcher, 1982; Hui & Triandis, 1985; Poortinga, 1989; Retief, 1988) have examined the problems of equivalence, comparability, validity, and bias, and offer useful discussions and alternative typologies to those of Sechrest et al. (1972) and Flaherty et al. (1988). One additional concept, metric equivalence, is of particular relevance to this paper and will be reviewed here.

2.3.8 Metric Equivalence

This has been achieved when the psychometric properties of the sets of data collected using the original and translated versions of the instrument exhibit essentially the same coherence or structure (Berry, 1980). Butcher and Pancheri (1976) discuss the importance of demonstrating equivalence of factor structure in assessing the adequacy of a translated instrument, while Ben-Porath (1990) regards failure to replicate factor structure as constituting a strong indication that a qualitative change may have occurred in the cross-cultural adaptation process. Metric equivalence should also be demonstrated at the item level (Clark, 1985). For example, item endorsement percentages should be comparable across the two versions of the instrument. Item response theory (Bontempo, 1993; Candell and Hulin, 1986; Hui and Triandis, 1985; Hulin, 1987) has been advanced as an alternative model for assessing this type of equivalence, and is reviewed later in this chapter.

As in the case of Casagrande's ends of translation, the practical translation of an assessment instrument will involve making compromises between the different types of equivalence. Werner and Campbell (1970) are critical of a tendency towards "a mischievous brand of logical positivism" which regards the original language instrument as the operational definition of the construct it is intended to measure. In research involving the translation of psychometric tests, this tendency may result in

greater emphasis being placed on vocabulary equivalence at the expense of other types such as idiomatic, experiential, and conceptual equivalence. In the words of Werner and Campbell (1970), "all of the translation effort has been directed toward representing it loyally in the target language. Figurative, metaphoric translation, in which target language idioms have been substituted for analogous idioms in the original, has been avoided due to fear of losing an item-by-item identity. Illustrative referents have been retained even though they represented familiar objects in the original language but exotic objects in the target language. Thus a fundamental asymmetry has resulted in which familiar, colloquial, accessible test items in English become exotic, awkward, and difficult items in the target language" (p. 414). By neglecting equivalence issues, such an approach will tend to exaggerate cross-cultural differences. Conversely, as suggested by Sechrest et al. (1972) and Butcher and Pancheri (1976) in their discussions of the *paradox of equivalence*, as more effort is invested in trying to achieve cross-cultural equivalence (e.g., via altering or eliminating nonequivalent items), the probability of finding cultural differences in the subsequent employment of the scale is decreased. Thus, in the extreme, translation for the sake of achieving equivalence may result in an instrument incapable of detecting the substantive cultural differences which a researcher may seek to study. Kleinman (1987) comments on a strong theoretical bias in psychiatric research towards finding cross-cultural similarities and "universals." In the absence of an ultimate criterion (Brislin et al., 1973), it is difficult to know when sufficient or optimal levels of equivalence have been achieved. Awareness of the above equivalence issues, judicious application of translation methods, close collaboration with members of the target culture, and a strategy of treating nonequivalence not simply as undesirable bias but as revealing cultural differences which should be analyzed, are seen as potential means of developing meaningful assessment tools.

2.4 MAXIMALLY TRANSLATABLE ENGLISH

Based on their experiences adapting instruments for cross-cultural use, Werner and Campbell (1970) and Brislin (1970, 1976, 1980, 1986; Brislin, et al., 1973) suggest guidelines intended to assist researchers in the design of instruments that should avoid undue difficulties in subsequent translation efforts. While these guidelines are written specifically to assist the writing of translatable *English*, similar rules may be applied to the construction of instruments in other languages, appropriate modifications to the rules being dependent on the specific languages concerned (Brislin, 1986; Retief, 1988). Brislin (1970) provides some empirical support for the effectiveness of such rules in reducing translation errors.

a) Use short, simple sentences of less than 16 words. There should be one dominant idea per sentence. Items involving more than one idea should be composed of a number of short sentences.

b) Employ the active rather than the passive voice. This will aid clarity by assisting translators in identifying the subject, object and verb in a sentence.

c) Repeat nouns instead of using pronouns. With more than one noun per sentence, the use of pronouns may result in unclear references due to vague noun-pronoun links.

d) Avoid metaphors and colloquialisms. These are unlikely to have direct equivalents in the target language and are consequently difficult to translate. As previously discussed, this is a controversial point.

e) Avoid the subjunctive (e.g., verb forms with could, would, or should). Other languages rarely have equivalent terms for the various forms of the English subjunctive.

f) Add sentences to provide context for key ideas, and reword key phrases to provide redundancy. Redundancy is intended to help translators (and respondents) to check meanings; if the meaning of an item is unclear in one phrase or sentence, it may be clarified in the light of another phrase or sentence. The longer the text, the less

it is open to ambiguous interpretation (Werner and Campbell, 1970).

g) Avoid adverbs and prepositions telling "where" or "when" (e.g., frequently, beyond, upper). There are often inadequate direct equivalents of these words in other languages. Further, the use of a term such as "frequently", for example, requires cross-cultural knowledge of base rates and subjective reactions to deviations from the base rates; what is regarded as frequent in one culture might not be regarded as such in another (Brislin, 1986).

h) Avoid possessive forms where possible. The concept of ownership as understood in Western cultures may differ markedly from that of other cultures.

i) Use specific rather than general forms (e.g., the specific animal, such as cows, chickens, or pigs, rather than the general term "livestock"). Different cultural groups do not necessarily categorise items in the same way. Consequently, any generalised label, if such exists in the target culture, is likely to differ to some extent to that of the source culture in the items which are subsumed by it.

j) Avoid words indicating vagueness (e.g., probably, frequently). It is best to specify a frequency in terms of a particular number of times per day, week, month, or year, as appropriate.

k) Use wording familiar to the translators where possible. Brislin (1986) asserts that "if wording is familiar to translators such that they can create a well-worded target language version, then that version will most likely be readily understandable to the eventual set of respondents in the data collection effort" (p. 148).

l) Avoid sentences containing two different verbs if the verbs suggest different actions. Confronted with two or more verbs in a sentence, translators may encounter difficulty attaching the relevant subject to the appropriate verb, thus further complicating the translation

process. In addition, the interpretation of subject's responses, in any language, is complicated in the case of such items; with multiple verbs indicating multiple actions, it is difficult to determine a respondent's reasons for endorsing or rejecting an item.

Detailed discussion of these points can be found in Brislin (1986). While these rules are intended to assist the development of new tests, they may also be of assistance in guiding the selection of existing instruments for cross-cultural adaptation, as well as in the modification of existing items to facilitate translation. The SCL-90-R, by comparison with many other instruments (e.g., the MMPI), conforms relatively well to these guidelines.

2.5 METHODS OF TRANSLATION

Several techniques or procedures aimed at achieving equivalence in cross-cultural translation are available and have been reported in the literature. Direct translation, described by Butcher and Pancheri (1976) as the most commonly used but least adequate translation procedure, disregards the issues of equivalence and makes no provision for assessing translation quality. In this procedure, a translator simply attempts to translate material from the source to the target language, no systematic effort being made to check the adequacy of such attempts. Due to the potential for translation errors and the inability of the direct translation method to detect such errors, Sechrest et al. (1972) contend that this method should be "rejected out of hand" (p. 51) in favour of more thorough alternatives. This section offers a critical review of the available techniques. Each of these techniques has associated strengths and weaknesses (Brislin, 1980) and address different aspects of equivalence. It is generally recommended that a combination of these techniques be used for any given translation

project, the weaknesses of each method being offset by the strengths of another.

2.5.1 Back-Translation

Back-translation is a commonly used procedure which has been widely recommended by cross-cultural researchers as a necessary, but not sufficient, step in the translation process. Sechrest et al. (1972) describe back-translation as a filter through which non-equivalent terms will not readily pass. According to this method, the original instrument is first translated into the target language by a bilingual translator (direct translation). This target language version is then independently (or "blindly") translated by a second bilingual back into the original language (back-translation). The researcher can then examine the back-translated version and compare it with the original, looking for discrepancies such as omissions, additions, substitution of terms, and ambiguous or nonsensical items. Using this method, even if the researcher does not know the target language, he or she will be in a position to make inferences regarding the quality of the target language translation, and thus exercise some control over the instrument development stage. If the two source language versions are considered to be equivalent, this would suggest that the target language version is equivalent to the source language forms. If discrepancies result they can be detected and further translation may be undertaken to achieve the required equivalence. However, as Brislin (1970) points out, back-translation is not foolproof and apparent equivalence may be spuriously created due to various factors. The following are examples of such factors (Brislin, 1970):

a) Translators may work according to unstated conventions for treating as equivalent certain words and phrases which are in fact not equivalent. Similarly, as discussed by Werner and Campbell (1970), errors may occur due to inaccurate but well-established translation-dictionary

equivalents which are used in both the forward and the back-translation.

b) A back-translator, faced with a poorly translated target language version, may be able to deduce what was intended and compensate accordingly when producing the back-translation. Thus a back-translator who is "overly competent" (Sperber et al., 1994) may give the misleading impression that the translation into the target language was satisfactory.

c) Grammatical structures of the source language may be retained incorrectly in the target language version. This version would consequently be relatively easy to back-translate, but would be "worthless" (Brislin, 1970, p. 186) for the purpose of studying target language monolinguals. Similarly, if translators are aware that their work is to be back-translated, they may use wording which they believe will back-translate well at the expense of wording which would be optimal for the target population (Geisinger, 1994).

It is also possible that an equivalent translation into the target language may have been achieved by the first translator, but that inaccurate back-translation may result in discrepancies between the original and back-translated versions. If the researcher relies on monolingual comparison of the original and back-translated versions, such discrepancies may be erroneously interpreted as stemming from an inadequate source to target language translation, thus confounding the translation effort.

It is important to recognise that while back-translation provides a useful quality check at the linguistic level (vocabulary equivalence) and some indication of experiential and conceptual equivalence (e.g., the lack of an equivalent term may indicate that the experience or concept is foreign to the target culture), it is not capable of reliably detecting nonequivalence at all levels. For example, nonequivalence may not be detected

in cases where the source language metaphors do not apply to the target language (idiomatic nonequivalence). Orley and Wing (1979, p. 518) provide an example from their adaptation of the Present State Examination (PSE) in Uganda. In the question, "Can people read your thoughts?", the verb "to read" was literally translated to the target language word meaning "to read a book" or "to study." This made no sense in the target language but could be back-translated into the original English.

In view of potential problems such as these, back-translation should not be regarded as sufficient to produce good quality, equivalent translations. The use of a committee of bilingual experts to assess translation efforts, and subsequent pretesting of the translated instrument are widely recommended addenda to the back-translation procedure. Brislin (1970) and Brislin et al. (1973) review various studies employing back-translation, and conclude that successful use of the method tends to be associated with the willingness of researchers to revise the source language version in order to facilitate a good final translation. This elaboration on the back-translation method is termed "decentering" (Werner & Campbell, 1970) and is discussed further below.

2.5.2 Decentering

As described by Werner and Campbell (1970), decentering requires that the source language version be open to revision, thereby seeking to avoid the ethnocentric imposition of instruments developed in one culture upon another. This involves "providing an understanding of cultural empathy, where content, context, and style may have to be adjusted to the receptor culture" (ibid., p. 399) in order to create a meaningful cultural equivalent. They go on to state that "in most social science applications of translation ... decentering and hence editing of the original schedule is unavoidable and desirable" (ibid. p. 407). Similarly, Sechrest et al. (1972) refer to decentering as

"an ultimate solution to the problem of translation" (p. 53) and view the process as a means of avoiding cultural and linguistic bias.

Decentering may be conducted after back-translation, modifications being made to items that are found to not translate well into the target language. Such modifications should be based on consultation with knowledgeable members of the target culture. The modified items may then be subjected to a further round of forward- and back-translation, and the results assessed again. Repeated rounds of this procedure may be employed until a satisfactory level of equivalence is achieved.

Brislin (1986) discusses decentering in relation to the emic-etic distinction: "If a concept 'survives' the decentering procedure, it is assumed to be *etic* since there must be readily available words and phrases in the two languages which the translators could use. If a concept is not in the final back-translated version, the reason could be that it is *emic*. That is, the concept might be readily expressible in only one of the languages" (p. 160, emphasis in original). In such cases, reasons why the items were not translatable should be investigated via discussion with the translators and bilingual consultants. Important cross-cultural differences which are uncovered in this way may be incorporated into the ongoing research process.

As discussed by John (1996), decentering may be seen as a prerequisite to the bilingual technique, discussed below, which involves administering both language versions of a test to a group of bilinguals in order to assess the equivalence of the two versions. By revising emic content (such as idiomatic expressions) in the source language instrument, the likelihood of response discrepancies across the two languages due to difficulty understanding the source language version is reduced.

Rogler (1989) cautions that the process of back-translation with decentering does not necessarily eliminate the problem of cultural bias. "Bias may still be present because the framework that structures the decentering process is derived from the research concepts expressed in the source language. Emic elements in the target culture may be left by the wayside" (ibid., p. 299). In the absence of a commitment to the examination and incorporation of indigenous manifestations of mental health problems, apparent cross-cultural homogeneity may be the result of the inappropriate application of culturally specific (emic) categories as if they were universally applicable (etic). Triandis (1980) refers to the latter approach as an "imposed etic". Sechrest et al. (1972) emphasise that decentering should go beyond the revision of source-language items in order to facilitate smooth translation, and should involve the construction of assessment instruments incorporating concepts generated jointly by both cultures, neither culture dictating to the other. Rogler (1989) asserts that the achievement of cultural sensitivity is an ongoing process that spans the entire research operation, and advocates that back-translation with decentering be used together with traditional ethnographic methods such as participant observation and interviews with members of the target culture.

Butcher (Butcher & Pancheri, 1976; Butcher & Clark, 1979) argues that in the case of existing instruments, particularly those which are widely used and have been standardised and extensively researched in the source culture, revision of the original test may not be a practical option. The benefits of revising would have to justify the costs, particularly the loss of norms and the accumulated body of research pertaining to the original instrument. There may also be copyright-related restrictions on the alteration of published tests. As an alternative, Butcher and Clark (1979) advocate "psychologically equivalent item substitutions" (p. 100), culturally-specific item content being replaced (in the target-language version only) with appropriate items which, based on anthropological data and

consultation with members of the target culture, are believed to be conceptually equivalent to the original items. The viability of this approach is challenged by Kortmann (1987) who comments on "the highly difficult task of, on the one hand adhering as closely as possible to the original text of the questions, in order to guarantee the instrument's standardization, while on the other hand having to take into account the norms and values of the culture in which the instrument is to be used. In consequence, a somewhat culture specific questionnaire will be produced, whose results can less easily be compared with 'the same' list, translated for another culture" (p. 568).

2.5.3 The Committee Approach

This approach involves a group of bilinguals working together, either to produce a translation or to evaluate and refine prior translation efforts. The main advantage of this approach is that errors and individual biases can be detected and rectified within the committee, the knowledge and expertise of individual members being pooled. Potential pitfalls include shared misconceptions of committee members and cultural or professional reluctance to criticise the work or opinions of others (Brislin et al., 1973).

2.5.4 The Bilingual Technique

The bilingual technique (Prince & Mombour, 1967) is a widely recommended psychometric means of assessing the equivalence of the original and translated versions of a test. The technique requires that a sample of bilingual subjects respond to both versions of the test, high positive correlations between original and translated items being regarded as indicative of equivalent translation having been achieved. Discrepant items (i.e., items which yield differing response frequencies across the two languages) can thus be identified, examined for sources of

nonequivalence, and subjected to attempts to improve translation. The procedure can be repeated until the desired equivalence is achieved.

A widely acknowledged disadvantage of employing bilingual subjects is that, being members of an atypical group, they may not be representative of the population for which the instrument is intended. Candell and Hulin (1986) make the point that the semantic structures of bilinguals may differ from those used by monolinguals when responding to the same items. For example, just as translators may draw on their knowledge of both languages in order to make sense of poor target language translations and compensate accordingly when producing back-translations (Brislin, 1970), bilingual respondents may similarly exploit their knowledge of the source language in order to infer what was intended from a poor target language translation. Translation equivalence may consequently be misevaluated by this technique.

A further potential confounding factor when employing the bilingual technique relates to the possibility that a bilingual will respond differently to the same item or stimulus depending on the language employed. In a study by Ervin (1964), for example, the TAT was administered in English and French on two different occasions to a sample of bilingual French subjects. Depending on the language used, the content of responses differed significantly and in ways that had been predicted on the basis of knowledge of the French and English cultures. Several more recent studies have found differences in bilinguals' responses to questionnaire items and interpreted the results in terms of social psychological explanations such as ethnic affirmation, cross-cultural accommodation, and social desirability. According to the ethnic affirmation hypothesis, testing in a language other than the subject's first language may intensify the individual's awareness of his or her ethnicity and increase the likelihood of responses which affirm that ethnicity (Bond & Yang, 1982; Yang & Bond, 1980). Conversely, cross-cultural accommodation refers

to the tendency to respond in a manner perceived to be appropriate to the culture associated with the language of testing. In other words, use of the second language may elicit values, attitudes, and role expectations of that culture which have been acquired in the process of learning the language. Based on the results of a study involving Chinese bilinguals, Bond and Yang (1982) argue that ethnic affirmation tends to occur on issues which are more central to the individual's cultural self-concept, while cross-cultural accommodation occurs on less important issues, on which he or she is prepared to compromise. Regarding social desirability, Marin, Triandis, Betancourt, and Kashima (1983) suggest that bilinguals may be more likely to respond in a socially desirable manner when answering in their second language. Tyson, Doctor, and Mentis (1988) suggest that such social desirability effects may be a function of the degree of status difference between the two cultures, a point that is of particular pertinence in the context of the present study, given South Africa's history of racial discrimination and apartheid.

Studies such as those of Yang and Bond (1980) and Bond and Yang (1982) provide grounds for questioning the assumption that bilinguals will respond in a similar manner to a given item regardless of the language employed. The suitability of the bilingual technique as a means of assessing translation accuracy is consequently in doubt. It should, however, be noted that subsequent studies have not supported their findings (Church, Katigbak, & Castaneda, 1988, Tyson et al. 1988), and that the studies may be criticised on methodological grounds, particularly failure to measure subjects' bilingual proficiency. This leaves open the possibility that the discrepant endorsement frequencies across languages may be due to respondents' differing levels of proficiency in the two languages. Another point is that, while these studies have tended to be relatively thorough in their translation procedures, using back-translation and decentering (Yang & Bond, 1980, Bond & Yang, 1982, Marin et al., 1983) and pretest procedures (Marin et al., 1983), these methods do not

guarantee equivalence. Brislin (1980) recommends that the bilingual technique be used in combination with these other methods in order to improve translation quality. It thus seems possible that item nonequivalence may have contributed to findings such as those of Yang and Bond (1980) and Bond and Yang (1982), and that these findings could even be interpreted as *supporting* the utility of the bilingual technique as a means of detecting nonequivalence. A final point is that studies investigating bilinguals' differential responses have tended to employ questionnaires dealing explicitly with attitudes and values and which were designed with the intention of emphasising cross-cultural differences (Bond & Yang, 1982; Yang & Bond, 1980). As stated by Bond and Yang (1982), "one can only affirm one's ethnicity on a given item if there is a difference between the perceived positions of one's group and another on that item" (p. 175). In the case of the SCL-90-R, which describes symptoms and which does not deal explicitly with cultural attitudes or values, effects such as cross-cultural accommodation, ethnic affirmation, and social desirability may expected to be of less relevance.

According to Brislin (1980), the advantage of the bilingual technique is its preciseness and potential for sophisticated statistical methods and concepts such as split-half reliability assessment. The technique also has advantages over the alternative technique of comparing item endorsement frequencies of independent monolingual samples from each culture, endorsement differences using the later method being potentially attributable to true cross-cultural differences in the distribution of the trait being measured (Hulin, 1987).

The bilingual technique has been seldom utilised in the South African context. A notable exception is John (1996) who employed the technique in assessing the fidelity of a Zulu language translation of the General Health Questionnaire (GHQ). Two problematic items were identified due to their low correlations across the two languages. In both cases, subsequent

analyses of the items suggested vocabulary nonequivalence as a possible source of the problem.

2.5.5 Pretesting

Prior to conducting the main study, all translated instruments should be field tested on a sample which is representative of the target population. Subjects may be asked to complete the instrument and then to explain their responses to each item. The purpose is to investigate the comprehensibility and acceptability of the instrument in the target population. Problems in this regard may be identified and attended to in the further revision of the instrument. Pretesting with monolingual subjects is seen as a means of avoiding, or compensating for, some of the problems associated with the reliance upon bilinguals in the translation process. Brislin et al. (1973) note that unforseen problems may be experienced by respondents despite the careful use of translation procedures such as back-translation. This is substantiated by Kortmann (1987) and Kortmann and Ten Horn (1988) who provide informative discussions of problems identified during the testing of an Amharic language translation of the Self Reporting Questionnaire (SRQ) in Ethiopia. Despite a careful translation procedure having been employed, problems, including linguistic and conceptual difficulties, resulted in forty percent of positive responses to the SRQ being rated as invalid. Language difficulties included items which were found to be too lengthy or complex for respondents to understand without clarification. Conceptual difficulties, which accounted for the majority of invalid endorsements, were identified by asking respondents to explain their answers as fully as possible and to clarify them with examples. For example, the item "Do you find it difficult to make decisions?" was often related to political circumstances and the respondents' limited social freedom, rather than the psychiatric intent of the item. Kortmann and Ten Horn (1988) conclude that "if the questionnaire is to be used in different cultures, an extensive study has to

be made of the meaning of the answers, which may lead to modifications in the way the question is formulated" (pp. 100-101).

Brislin (1980) and Retief (1988) emphasise that the efficacy of pretesting depends both on the quality of the investigation and the extent to which obtained feedback is utilised in revising the test.

2.5.6 Item Response Theory

Item response theory (IRT) is a statistical approach which has been applied to the assessment of the translation fidelity of psychological scales (Bontempo, 1993; Candell and Hulin, 1986; Hui and Triandis, 1985; Hulin, 1987). This approach has been advanced by psychometricians as an alternative to classical measurement models which are argued to have serious limitations and which are expressed at test level rather than individual item level (Hambleton, Swaminathan, and Rogers, 1991). IRT is based on the proposition that the probability of responding to a given item in a specified manner is a function of an individual's status on the underlying "latent trait" which is being measured by the test. The term "latent trait" refers to a statistical construct, and is often estimated from the total score obtained on the test being assessed (Anastasi, 1990). These probabilities are represented by item characteristic curves (ICCs). Statistical tests may be used to examine differences between ICCs obtained from different cultures. Individuals from different cultures (and who speak different languages) who possess the same amount of the latent trait should respond in a specified manner with the same probability to an item or its equivalent translation (Candell and Hulin, 1986). Identified differences may indicate lack of equivalence of a particular item while similar ICCs across cultures may be regarded as indicating psychometric equivalence of the source and translated item (Hui and Triandis, 1985, Hulin, 1987). Biased items, whether the

result of poor quality translation or emic content, can thus be identified and subjected to further efforts to achieve equivalence.

IRT models rest on a set of assumptions regarding the characteristics of the data to which the models are applied. These assumptions include the item pool being unidimensional and that the particular model chosen adequately fits the data. A discussion of IRT assumptions can be found in Hambleton et al. (1991). The assumption of unidimensionality will be discussed here in order to illustrate potential shortcomings of applying IRT to research such as the current study.

The assumption of unidimensionality, that is, that a single latent trait is measured by the items that make up the test, must be met in order for item parameters to be validly estimated. A number of problems relating to this assumption should be noted. Firstly, the unidimensionality assumption is contrary to the assumptions underlying Western psychiatric nosology as expressed in the DSM-IV (American Psychiatric Association, 1994) and multidimensional assessment instruments such as the SCL-90-R (Derogatis, 1983). As Bontempo (1993) asserts, no psychological instrument is likely to be strictly unidimensional. It may, however, be possible to accommodate the unidimensionality assumption by separately analyzing each subscale of multidimensional tests (Candell & Hulin, 1986). Secondly, the assumption cannot be strictly met due to the influence on test performance of variables such as levels of motivation, social desirability, and test anxiety. Nevertheless, proponents of IRT assert that the unidimensionality assumption will be "met adequately" by the presence in a set of test data of a *dominant* factor that influences performance on the test (Hambleton et al., 1991). However, as Bontempo (1993) points out, no explicit criterion exists for assessing whether an item pool is sufficiently unidimensional to permit IRT analysis, and according to Anastasi (1990), due to the likelihood of curvilinear item

intercorrelations, unidimensionality cannot be demonstrated by the usual (linear) factor-analytic procedures.

The unidimensionality assumption thus poses a problem to the valid application of IRT models to real psychological data. Multidimensional models do exist but, according to Hambleton et al. (1991), are more complex and have yet to be sufficiently developed and tested. Consequently, IRT methods will not be utilised in the current research.

CHAPTER 3

THE SCL-90-R

3.1 INTRODUCTION

This chapter describes and critically discusses the SCL-90-R, and reviews the research literature relating to the instrument's reliability, validity, and factorial structure.

The SCL-90-R (Derogatis, 1983) is a widely used multidimensional, 90-item, self-report symptom inventory designed to measure symptomatic psychological distress. The inventory was developed from the Hopkins Symptom Checklist (HSCL; Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974), with the prototype version (SCL-90) being published in 1973 (Derogatis, Lipman, & Covi, 1973). The copyrighted SCL-90-R is almost identical to the SCL-90, with two items (both on the Anxiety dimension) being replaced and a further two items being slightly modified.

Psychological distress is measured in terms of nine primary symptom dimensions and three global indices of distress. The symptom dimensions represent the following constructs:

I. Somatization (SOM): reflects distress arising from perceptions of bodily dysfunction;

II. Obsessive-Compulsive (O-C): reflects symptoms that are closely identified with the clinical syndrome of the same name;

III. Interpersonal Sensitivity (INT): focuses on feelings of personal inadequacy and inferiority, particularly in comparisons with others;

IV. Depression (DEP): reflects a broad range of cognitive and somatic correlates of clinical depression;

V. Anxiety (ANX): reflects symptoms associated clinically with manifest anxiety;
VI. Hostility (HOS): reflects thoughts, feelings, and behaviours characteristic of angry affect;
VII. Phobic Anxiety (PHOB): focuses on the more pathognomic and disruptive manifestations of phobic behaviour, particularly agoraphobic symptoms;
VIII. Paranoid Ideation (PAR): represents paranoid thinking characterised by projective thought, hostility, suspiciousness, grandiosity, centrality, fear of loss of autonomy, and delusions; and
IX. Psychoticism (PSY): represents a range of severity, from interpersonal alienation to first-rank symptoms of schizophrenia.

Each of these dimensions consists of between 6 and 13 items. The score for each dimension is calculated as the mean of its constituent items.

More detailed clinical definitions of each of the SCL-90-R symptom dimensions are provided by Derogatis (1983). These dimensions were established by means of rational selection of symptom clusters, with subsequent empirical verification via confirmatory factor analysis (Derogatis & Cleary, 1977b). Studies relating to the factor structure of the SCL-90-R are reviewed below.

The three global indices of distress (Global Severity Index (GSI), Positive Symptom Distress Index (PSDI), and Positive Symptom Total (PST)), reflect related but distinct aspects of general psychological distress (Derogatis, Yevzeroff, & Wittelsberger, 1975). The GSI combines information on the number of symptoms and the intensity of distress; the PSDI is a measure of intensity only; the PST is a count of the number of symptoms only. According to Derogatis (1983), the GSI is the best single indicator of distress, and should be used in most cases where a single summary score is needed.

Interpretation of the SCL-90-R thus involves three levels: the global scores indicate the degree of general psychological distress; the primary symptom dimensions provide a profile in terms of specific areas of psychopathology; and individual items provide information at the level of discrete symptoms.

Each item is rated on a 5-point scale according to how much the symptom has distressed the respondent over a specified period of time, usually the preceding 7 days (0 = "not at all", 1 = "a little bit", 2 = "moderately", 3 = "quite a bit", 4 = "extremely"). The results of research into the effects of the number of scale points on the psychometric properties of psychological scales are inconsistent. There is no ultimate criterion for determining the optimal number of scale points to use for psychometric testing, this being dependent on the inherent discriminability of the particular stimuli being rated (Garner, 1960) as well as the respondent's discriminative capabilities and willingness to utilise such capabilities in the testing context (Guilford, 1954). A Monte Carlo study by Lissitz and Green (1975) provides support for the reliability of 5-point scales, while Velicer and Stevensen (1978), working with the Eysenck Personality Inventory, found that a multi-choice response format yielded a superior factor structure to that of the standard two-choice format. Correspondingly, Comrey (1988) discusses some of the psychometric difficulties associated with dichotomous data and advocates a minimum of five numerical response categories. He asserts that multi-choice item formats produce better scales which are more reliable, contain more information, and are less subject to statistical distortions than two-choice formats. As Garner (1960) points out "it is clear that information transmission cannot be lost by increasing the number of rating categories. Therefore, it is better to err on the side of having too many categories than to err by having too few" (p. 352).

In the South African context, despite the rationale outlined above, Mauer and Retief (1985) and Taylor and Boeyens

(1991) advocate a two-choice format as reducing cross-cultural stylistic differences which may negatively affect the metric properties of psychological instruments. Similarly, John (1996), in his translation of the General Health Questionnaire into Zulu, opted for the simplified Yes/No (*Yebo/Cha* in Zulu) over the original four-point GHQ response scale (ranging from "not at all" to "much more than usual") since "the present teams of translators were in agreement that the original response format of the GHQ was particularly difficult to convey in Zulu and would ultimately be confusing for the respondents" (p.81). The decision to utilise the simplified format was thus made in the absence of empirical support and without having conducted a pretest study using the multi-choice format. Conversely, Michelson (1991), in her study involving a Zulu translation of the Brief Symptom Inventory (BSI; Derogatis & Melisaratos, 1983), a brief (53-item) form of the SCL-90-R, utilised the original 5-point format (which is identical to that of the SCL-90-R) and reported no difficulties associated with the use of the instrument with a Zulu population. Since the SCL-90-R is designed to measure not only the presence of symptoms but the *intensity* of distress experienced as a result of the symptoms, it seems appropriate to retain the original response format unless this is clearly shown (e.g., in pretest studies) to be inappropriate for use with the target population.

While the global measures (GSI, PSDI, PST) provide information regarding individuals' response styles, specifically regarding the tendency to repress ("fake good") or augment ("fake bad"), thus assisting in the identification of invalid protocols (Derogatis, 1983), formal validity scales to detect response distortions are not built into the SCL-90-R. Derogatis's (1983) explanation of this omission includes the difficulty, in the case of pure symptom inventories (unlike, for example, personality inventories), of constructing suitable validity scale items that are not easily discernable from substantive items. Derogatis (1983) also questions whether response sets have substantial systematic effects on clinical

self-report inventories, and argues that validity scales "do not result in a reliable and generalizable adjustment of scores for systematic response distortions across populations" (p. 35).

The judiciousness of omitting validity scales from the SCL-90-R has been challenged by studies conducted by Dinning and Evans (1977) and Brophy, Norvell, and Kiluk (1988). These studies found significant correlations between the SCL-90-R dimensions and the validity scales of the MMPI. Working with an inpatient psychiatric sample, Dinning and Evans (1977) found that the SCL-90 dimensions correlated significantly and consistently with the L and K scales of the MMPI, indicating that defensive and dissimulating patients tended to obtain lower scores on all nine dimensions of the SCL-90. In addition, the SCL-90 dimensions, with the exceptions of Depression and Hostility, were significantly correlated with the MMPI F scale, indicating that patients with a "fake bad" response set tended to obtain higher scores on most SCL-90 dimensions. These findings are substantiated by Brophy et al. (1988), who found significant correlations between all nine SCL-90-R dimensions and the three validity scales of the MMPI in an outpatient sample. Brophy et al. (1988) conclude that caution should be exercised in the interpretation of SCL-90-R scores, and suggest that "the addition of a validity scale might help to correct this deficit and improve clinical usefulness" (p. 339). They do not, however, suggest items which may be suitable for a SCL-90-R validity scale. In another study relating to the validity of responses, Lees-Haley (1989) reports that untrained student volunteers were able to simulate psychopathology on the SCL-90-R, and advises caution in the application of the instrument, particularly in settings where deliberate manipulation of responses may be an issue.

In developing the test items, attention was given to using the most basic level of vocabulary possible (Derogatis et al., 1973). Examination of the items reveals that they are grammatically and syntactically quite simple, and conform

relatively well to Brislin's (1986) guidelines for writing maximally translatable English. While the SCL-90-R has been translated into at least 24 languages (Derogatis & Lazarus, 1994), according to the publishers (NCS Assessments, personal communication, 1996) there is currently no Zulu-language version of the test. However, a Zulu translation of the 53-item BSI was produced by Michelson (1991) and used in a study of a Zulu-speaking population in the Pietermaritzburg area of KwaZulu-Natal. While the back-translation method was utilised in the study, no detail regarding the translation process is provided by Michelson (1991), and no effort was made to establish the validity of the translated instrument as applied to the target population.

Several studies employing various translations of the SCL-90-R have been published in the major journals. These languages include Dutch (Hafkenscheid, 1992; Koeter, 1992), German (Rief & Fichter, 1992), Hebrew (Roskin & Dasberg, 1983; Schwartzwald, Weisenberg, & Solomon, 1991), Hmong (Westermeyer, Vang, & Neider, 1983), Korean (Noh & Avison, 1992), and Portuguese (Simoes & Binder, 1980). Unfortunately, despite the recommendations of researchers such as Brislin (1970) and Butcher and Clark (1979) for detailed reporting of the translation process, these articles provide very little such information.

Administration of the test is simple and may be performed by nonprofessionals (Derogatis, 1983), thereby increasing its potential utility as a screening instrument in the South African primary health care context. Verbal administration of the SCL-90-R is possible, another important consideration in the context of the current study, given the high levels of illiteracy in South Africa. Derogatis (1983) reports that "narrative report" administrations in the United States, in medical settings where debilitation or physical trauma prevented patients from physically completing the test, "did not reveal any consistent biases associated with the technique" (p. 4).

Similarly, Westermeyer, Vang, and Neider (1983), utilising a Hmong translation of the SCL-90 with a sample of refugees in the United States, report no significant differences between the responses of subjects who filled out their own forms and illiterate subjects who had the inventory read to them.

The SCL-90-R was developed for use with a broad range of respondents. Formal published norms, based on samples in the United States, are available for community nonpatients, psychiatric outpatients, psychiatric inpatients, and adolescent nonpatients (Derogatis, 1983). Separate norms are provided for males and females in each of the above categories. Geriatric and other specialised norms are being developed in the United States (Derogatis & DellaPietra, 1994). In addition to the formal norms, SCL-90-R profile data are available for a variety of specific clinical groups (Steer, 1982).

3.2 RELIABILITY

Internal consistency and test-retest reliability data for the SCL-90-R symptom dimensions are reported in Derogatis (1983). Internal consistency coefficients (coefficient alpha), based on the responses of 219 "symptomatic volunteers", are reported as satisfactory, ranging from a low of .77 for psychoticism to a high of .90 for depression (average = .84). Similarly substantial internal consistency coefficients have been reported in other studies (Horowitz, Rosenberg, Baer, Ureno, & Villasenor, 1988; Mazmanian, Mendonca, Holden, & Dufton, 1987).

Test-retest coefficients, based on a sample of heterogeneous psychiatric outpatients with one week elapsed between testings, are also satisfactory, ranging from .78 for hostility to .90 on the phobic anxiety dimension (Derogatis, 1983). Horowitz et al. (1988) investigated the test-retest reliability of the SCL-90-R with a psychiatric outpatient sample

over 10 weeks, and reported subscale coefficients ranging from .70 for obsessive-compulsive to .83 for paranoid ideation, and a GSI coefficient of .84. These figures are particularly satisfactory considering the inherent instability of psychological symptomatology relative to, for example, traits such as those measured by intelligence tests or personality inventories.

3.3 VALIDITY

The SCL-90-R has been utilised in a variety of contexts and has been found to be sensitive to psychological distress arising from a broad variety of conditions, including depression (Derogatis & Wise, 1989; Prusoff, Weissman, Klerman, & Rounsaville, 1980; Wetzler, Kahn, Strauman, & Dubro, 1989), anxiety disorders (Cameron, Thyer, Nesse, & Curtis, 1986; Derogatis & Wise, 1989; Horowitz, Wilner, Kaltreider, & Alvarez, 1980; Rief & Fichter, 1992), sexual dysfunctions (Derogatis, Meyer, & King, 1981), eating disorders (Peveler & Fairburn, 1990), substance abuse (Platt, Steer, Ranieri, & Metzger, 1989), chronic pain (Jamison, Rock, & Parris, 1988; Peltz & Merskey, 1982), and toxic exposure (Uzzell & Oler, 1986).

In addition, the SCL-90-R has been shown to be sensitive to change due to a variety of therapeutic modalities, including psychotropic medications (Ballenger, Burrows, Dupont, Lesser, Noyes, Pecknold, Rifkin, & Swinson, 1988; Ravaris, Robinson, Ives, Nies, & Bartlett, 1980), psychotherapeutic interventions (Shapiro & Firth, 1987; Woody, Luborsky, McLellan, O'Brien, Beck, Blain, Herman, & Hole, 1983), and relaxation techniques (Carrington, Collings, Benson, Robinson, Wood, Lehrer, Woolfolk, & Cole, 1980). Due to its high test-retest reliability and purported lack of significant "practice" effects (Derogatis, 1983), the SCL-90-R is particularly useful for repeated assessments across time.

In the context of neuropsychological testing, Lezak (1995) reports that the SCL-90-R has been used with a variety of neuropsychological disorders, including stroke, head trauma, multiple sclerosis, AIDS, and exposure to neurotoxins. Lezak (1995) recommends the SCL-90-R as being particularly useful in identifying patients with attentional and memory disorders. Such patients "tend to have score elevations particularly on the Obsessive-Compulsive scale, as they check items having to do with mental inefficiency (problems in concentrating, drawing a mental blank), poor memory, and with techniques to compensate for these problems, such as working slowly to guard against errors, or double-checking their work" (ibid., p.789). This is consistent with the work of O'Donnell, DeSoto, and Reynolds (1984) who suggest an 8-item "Cognitive Deficit" subscale of the SCL-90-R, consisting predominantly of items from the Obsessive-Compulsive subscale, which they found to distinguish patients with compromised neuropsychological functioning ($N = 50$; evaluated with the Halstead-Reitan Battery) from non-patients and psychiatric outpatients according to the normative data provided in the SCL-90-R manual (Derogatis, 1983).

Derogatis, Rickels, and Rock (1976) reported a high degree of convergent and discriminant validity for the SCL-90 in a study which compared its symptom dimensions with those of the MMPI in a group of 209 symptomatic volunteers for psychotherapeutic drug trials. Each of the SCL-90 dimensions correlated most highly ($r > .40$) with a comparable MMPI scale (except for the obsessive-compulsive dimension, for which there is no directly comparable MMPI scale), while having lower correlations ($r < .40$) with noncomparable MMPI scales. A British study by Wilson, Taylor, and Robertson (1985) provides additional evidence of convergent validity, reporting satisfactory correlations between the SCL-90 subscales and corresponding scales of the Present State Examination (PSE) and the Comprehensive Psychopathological Rating Scale (CPRS). Dinning and Evans (1977) compared psychiatric inpatients' responses on the SCL-90 to those on the Beck Depression

Inventory (BDI), the State-Trait Anxiety Inventory (STAI), the Whitaker Index of Schizophrenic Thinking (WIST), and the MMPI. They reported moderate to high convergent validity for the SCL-90 dimensions but poor discriminant validity, since the dimensions correlated significantly with virtually all other measures included in the study. The findings of Dinning and Evans (1977) are supported by Brophy et al. (1988) who found reasonable convergent validity but poor discriminant validity of the SCL-90-R subscales, in relation to the MMPI and the BDI, in an outpatient population. The results of Dinning and Evans (1977) and Brophy et al. (1988) question the utility of the SCL-90-R as a multidimensional measure of psychopathology and suggest that the instrument may be best conceptualised as a measure of general psychological distress. This issue is discussed further in the context of factor analytic studies of the SCL-90-R.

3.4 FACTOR STRUCTURE

The multidimensional structure of the SCL-90-R was developed by means of clinical-rational methods to describe the symptomatology of patients in terms of nine primary symptom dimensions (Derogatis, 1983). In an effort to provide support for the construct validity of the instrument, Derogatis and Cleary (1977b) compared this hypothetical structure with an empirically-derived dimensional structure based on a sample of 1,002 heterogeneous psychiatric outpatients in the United States. They reported satisfactory hypothetical-empirical matches for eight of the nine dimensions. The ninth dimension, Psychoticism, while producing "good" agreement in the Procrustes analysis, performed less well in terms of the varimax procedure, causing Derogatis and Cleary (1977b) to conclude that "more work may need to be done on the definition of this construct [with the focus on] alteration rather than abandonment" (p. 989). Despite the stated intention, to date no alteration of the Psychoticism dimension has been conducted. Derogatis and Cleary

(1977b) interpret their results as providing substantive support for the hypothesised dimensional structure of the SCL-90.

Based on the same sample of outpatients, Derogatis and Cleary (1977a) report "substantial" factorial invariance across gender for eight of the primary symptom dimensions and "moderate" levels of invariance for the ninth (Paranoid Ideation). They interpret these results as leaving "little doubt as to the generalizability of these symptom constructs across the variable of gender, and coupled with previous work on a number of the dimensions, suggest that the domain of psychopathology as measured by the SCL-90 may prove to be a highly consistent one" (ibid., p. 354).

Subsequent research has questioned the dimensional structure of the SCL-90-R, as well as its purported invariance across different groups. Hoffmann and Overall (1978) analyzed the responses of 358 psychiatric outpatients by means of the varimax method and obtained only five interpretable factors, labelled Depression, Somatization, Phobic Anxiety, Functional Impairment, and Hostile Suspiciousness. In addition, the large proportion of the total variance accounted for by the first unrotated factor (6.45 times more than the next largest factor) and high intercorrelations among the factors suggest that the instrument is a unitary measure of global discomfort rather than a measure of distinct dimensions of psychopathology as proposed by Derogatis (1983).

Also working with a psychiatric outpatient sample (N = 327), Evenson, Holland, Mehta, and Yasin (1980) obtained a twelve factor solution which differed substantially from that proposed by Derogatis (1983). Most notably, no independent Anxiety dimension was found. Despite the fact that the first unrotated factor ("Agitated Depression") accounted for nine times the variance of the second largest factor, thus corroborating the findings of Hoffmann and Overall (1978),

Evenson et al. (1980) recommend that the SCL factor dimensions be retained subject to further research.

An instructive paper by Cyr and Atkinson (1986) criticises Evenson et al. (1980) on several methodological points, including a low subject-to-variable ratio (3.6 to 1)[1], low numbers of items with significant loadings on a factor (several factors consisted of only two or three items), and a high frequency of items which load significantly on more than one factor. Reanalysing the Evenson et al. (1980) data according to more stringent criteria, Cyr and Atkinson (1986) suggest that only five factors should have been interpreted.

Holcomb, Adams, and Ponder (1983) investigated the factor structure of the SCL-90 with a sample of 451 inpatients and obtained a nine factor solution. These factors differ somewhat from those of Derogatis (1983; Derogatis and Cleary, 1977b) with no clearly identifiable factors representing Interpersonal Sensitivity, Anxiety, or Psychoticism. The first factor, "Depression", accounted for 8.8 times the variance of the next largest factor, further substantiating previous findings suggesting that the SCL-90-R may be a unidimensional measure of general distress.

In a recent study, Carpenter and Hittner (1995) assessed the dimensional properties of the SCL-90-R for male (n = 402) and female (n = 227) inpatients diagnosed with concurrent substance abuse and other Axis I disorders. Exploratory factor analyses yielded a four-factor structure for males and a seven-factor structure for females. These results thus contradict the major findings of Derogatis and Cleary (1977a, 1977b), failing

---

[1]The smaller the subject-to-variable ratio, the greater the possibility of obtaining spurious factors due to chance (Nunnally, 1978). Comrey (1988) recommends a ratio of five subjects per variable, which would require a sample of 450 subjects in the case of the SCL-90-R. Brislin, Lonner, and Thorndike (1973) are more stringent, maintaining that "a bare minimum sample should include 10 times as many subjects as variables" (p.279).

to support the postulated factor structure and indicating gender differences in the number and composition of symptom dimensions. For both genders, the first factor (labelled "general psychopathology") accounted for more than six times the proportion of variance accounted for by the second factor, once again suggesting that the SCL-90-R is a unidimensional measure of psychological distress. Further support for the unidimensional status of the SCL-90-R is provided by Bonynge (1993), Brophy et al. (1988), Clark and Friedman (1983), Cyr, Doxey, and Vigna (1988), Hafkenscheid (1993), Mazmanian et al. (1987), Schwartzwald et al. (1991), Steer, Clark, and Ranieri (1994), and Strauman and Wetzler (1992). Thus the majority of factor analytic studies of the SCL-90-R suggest that the instrument would best be considered as a measure of global distress and that discretion should be exercised in the interpretation of primary symptom dimensions.

The unidimensionality of the SCL-90-R is not accepted by all researchers. Rief and Fichter (1992) analyzed the responses of 899 psychiatric inpatients to a German translation of the SCL-90-R and obtained a nine factor solution which differed substantially from that of the original version. Based on their results as well as those of previous studies, including Hoffmann and Overall (1978) and Holcomb et al. (1983), Rief and Fichter (1992) suggest that the original version of the SCL-90-R should be revised to include a "Sleep Disturbance" factor, while the Anxiety and Phobic Anxiety factors should be combined and the Paranoid Ideation and Psychoticism factors should be reformulated. Discriminant analyses using the original and proposed new factorial structures produced average hit rates of 67% for the original and 74% for the new factorial structure in distinguishing between patients with dysthymia, anxiety disorders and anorexia nervosa. Rief and Fichter (1992) assert that the SCL-90-R's satisfactory specificity in describing dysthymia, anxiety disorders, and anorexia nervosa support its multidimensional utility, and, despite the findings of previous

studies, reject the argument that the inventory is primarily a measure of general distress.

Schwartzwald et al. (1991) studied the responses of two groups of Israeli soldiers to a Hebrew version of the SCL-90-R. The one group consisted of soldiers identified as suffering from combat stress reaction, while the other group was regarded as healthy. While a Procrustes analysis supported the original structure proposed by Derogatis and Cleary (1977b), subsequent varimax analysis did not support the original structure and yielded different factor structures for the two groups. Schwarzwald et al. (1991) interpret their results as suggesting that "the SCL-90-R may be either a measure of general discomfort or, more likely, that different factors may be needed for different populations" (p. 385).

Finally, in a study which assessed the performance of the SCL among different ethnic groups, Takeuchi, Kuo, Kim, and Leaf (1989), analyzed the responses of four ethnic groups in Hawaii (Caucasians, Filipinos, Japanese, and Native Hawaiians) to 54 items of the SCL-90, which form five of the nine scales (Depression, Somatization, Anxiety, Interpersonal Sensitivity, and Obsessive-Compulsive). Using a Procrustes rotation they found that the scale item loadings generally did not correspond to the hypothesised factors. The Caucasian sample produced a better fit to the hypothesised factors than the other groups, with the Native Hawaiians producing the worst fit. These results emphasise the need to assess, rather than assume, the validity of psychological scales as applied to different cultural groups which may vary in their experience and reporting of symptoms.

To summarise, a number of factor analytic studies of the SCL-90-R have been conducted, using various methods of factor analysis and with samples from a variety of clinical populations, nationalities, and ethnic groups. A review of these studies suggests:

(a) the original nine-dimensional structure proposed by Derogatis (1983) lacks empirical support;
(b) the instrument is probably best conceptualised and utilised as a unidimensional measure of general psychological distress, although judicious use of subscale scores may facilitate clinical interpretation; and
(c) the factor structure of the instrument should be empirically established for each population with which it is to be used.

CHAPTER 4

TRANSLATION OF THE SCL-90-R INTO ZULU

4.1 THE TRANSLATION TEAM

In addition to bilingual fluency, it is recommended that translators involved in the cross-cultural adaptation of assessment instruments be knowledgeable about both the source and target cultures, and familiar with the content area of the instrument (Brislin, 1970; Butcher & Clark, 1979; Geisinger, 1994). The author was cognizant of the above criteria in recruiting the translators and consultants who participated in the current study.

The team of translators and bilingual consultants consisted of three psychologists, a psychiatric nurse, and three experts in the Zulu language. Six of the team members were native Zulu speakers and one member a native English speaker. All were fluent in both Zulu and English. The psychiatric nurse had several years experience working with Zulu patients in both psychiatric and medical hospital settings. The language experts, who included the native English speaker, were lecturers at the Zulu Department of the University of Natal at Pietermaritzburg, and all were experienced translators. The translators were contracted on a favour basis.

4.2 INSTRUCTIONS TO TRANSLATORS

The purpose of the study and intended uses of the translated checklist were discussed with each translator, and specific guidelines were provided for the translation of the SCL-90-R into Zulu (see Appendix A). These guidelines are based on the recommendations of Sechrest et al. (1972) regarding

translation equivalence. The translators were also instructed to rate each item according to the difficulty encountered in achieving what they regarded as an equivalent translation (1 = relatively easy to translate, equivalent meaning achieved; 2 = some difficulty in translation, equivalence may be questionable; 3 = considerable difficulty in translation, equivalence is questionable). Translators were further requested to make brief notes of particular difficulties encountered and the decisions they made in attempting to resolve these difficulties. These ratings and notes were intended to assist the researcher in identifying SCL-90-R items which may be problematic due, for example, to emic content. In addition, all translators were asked to comment on whether they regarded the checklist as culturally appropriate for Zulu people, and to indicate particular items which they believed may be culturally irrelevant, unacceptable, or otherwise inappropriate.

4.3 THE TRANSLATION PROCESS

4.3.1 First Round of Back-Translation, and Decentering

The original English version of the SCL-90-R was translated into Zulu by one of the psychologists. Thereafter, a second psychologist, who was not familiar with the SCL-90-R, was given the Zulu translation and asked to translate it into English. These two steps resulted in an initial Zulu translation and an English back-translation which could be compared with the original SCL-90-R.

The two translators, the researcher, and an additional bilingual psychologist formed a committee to assess the initial translation efforts. The purpose of this step was to identify items which were difficult to translate into Zulu and, where necessary, to revise the original SCL-90-R items in order to facilitate subsequent translation.

All 90 items, as well as the instructions, response categories, and example, were reviewed. Particular attention was paid to items which had (a) been rated by the translators as presenting problems in achieving equivalent translation (ratings of 2 or 3), and/or (b) resulted in back-translations which were identified by the researcher as being nonequivalent to the original English items.

Back-translation is not foolproof (Brislin, 1970), and discrepancies between the original and back-translated versions may be due to two possible sources of error. First, the translation into Zulu may have been inadequate, the back-translation performing its function of assisting in the identification of errors. Second, the translation into Zulu may have been adequate but the back-translation inadequate, thus misleadingly suggesting that the Zulu translation was incorrect. Conversely, as discussed by Brislin (1970), back-translation may serve to disguise translation errors. For example, the back-translator may deduce what was intended from a poorly translated target translation and compensate accordingly. Consequently, while the back-translation provided a valuable guide in the identification of translation difficulties, it was necessary to review and discuss each item within the committee regardless of the degree of equivalence between the original and back-translated English versions.

Six items (numbers 30, 47, 57, 64, 65, and 86) were identified by the committee as requiring decentering in order to facilitate translation into Zulu. Two of these items were problematic due to the use of idiomatic expressions in the original English, three due to experiential inequivalence, and one due to grammatical-syntactical inequivalence. Examples of these types of problems and how they were resolved within the committee are discussed below:

a) Idiomatic inequivalence

Item number 30, which forms part of the depression dimension, refers to the experience of "feeling blue". While in the English language colours are associated with certain emotional states (see Dunnigan et al., 1993), this is not the case in Zulu and such an item is meaningless if literally translated. Based on committee discussions, this original item was revised to refer to the experience of feeling "sad or depressed". This was regarded as preserving the psychiatric intent of the item while being relatively easy to translate into Zulu.

b) Experiential inequivalence

Item number 47, which forms part of the phobic anxiety dimension, refers to anxiety associated with travelling on "buses, subways, or trains". Since subway travel is experientially foreign to the target population of this study, the committee decided to delete the word "subways", the revised item referring to "buses or trains". It was also discussed in the committee that the Zulu term for "buses", *amabhasi*, would tend to be interpreted by respondents as referring to minibus taxis, a common form of transport, as well as larger buses. It was believed that the revised item would be relevant to the cultural context. There was, however, some question as to the criterion equivalence of the item. This pertained to the questionable face validity of the item as an index of psychopathology in the target population; given South Africa's history of violence related to buses, taxis, and trains, a degree of anxiety associated with utilising these modes of transport may reflect well-founded concerns rather than phobic anxiety. Nevertheless, the revised item was accepted by the committee, subject to empirical validation.

c) Grammatical-syntactical inequivalence

Item number 86, which forms part of the anxiety dimension, refers to distress due to "thoughts and images of a frightening nature". This item was revised to read "frightening thoughts

and images", since the wording of the original item was regarded as clumsy and potentially confusing. For example, it is was discussed that a translator or second-language respondent may misinterpret the item as referring to fears relating to the natural environment ("nature").

4.3.2 Second Round of Back-Translation, and Selection of the Best Zulu Translation

The decentered English language version was given to two translators from the Zulu Department to translate into Zulu. This pair included the native English speaker. These translators were accustomed to working on translation tasks as a team, and agreed to participate in the study on condition that they work together.

This second Zulu translation was given to another translator, also a member of the Zulu department, for independent translation back into English. Upon completion of this step, two rounds of translation and back-translation had been completed, resulting in two Zulu versions and two English back-translations.

A committee, consisting of the author, the three Zulu psychologists, and the psychiatric nurse, met to consider the two Zulu translations in order to arrive at a single satisfactory Zulu version of the SCL-90-R.

Each item was considered individually in the committee, the pros and cons of different Zulu versions of the items being openly debated. All members of the committee participated enthusiastically in these discussions, drawing on their extensive knowledge of language, culture, and psychopathology. There was no apparent evidence of any cultural or professional reluctance to criticise the work or opinions of others (cf., Brislin et al., 1973). The thoroughness with which the

committee executed its task is reflected in the fact that these discussions required several meetings, totalling approximately eight hours of committee time.

Of the 91 items (including the example), eleven were identical across the two Zulu translations. Of the remaining 80 items, the majority could be described as showing reasonably close correspondence, with only slight variations in grammar or choice of words by the different translators. Bearing in mind Werner and Campbell's (1970) view of translation as a form of paraphrase, each item having more than one possible suitable translation, it was nevertheless necessary to select a single "best" version of the Zulu SCL-90-R.

The guiding criteria for the selection of the best Zulu translation of each item were as follows:

a) the Zulu item should most closely reflect the meaning of the original English item; and

b) the item should read naturally, be as simple as possible, and likely to be understood by the widest range of Zulu-speaking individuals from different geographical areas and of different educational backgrounds.

In the case of 48 of the items the first translation was selected, while for 25 items the second translation was regarded as superior. In the case of 7 items it was decided that a combination of elements from both translations was most desirable. These items were rewritten within the committee.

A recurring issue throughout the committee discussions related to the specificity of terms in the Zulu and English languages. While the English language contains a rich variety of terms referring to different symptoms and emotional states, in many cases specific Zulu equivalents do not exist. The Zulu terms tend to be less specific, so that a single term may be interpreted as referring to different states (e.g., anxiety or depression) depending on the context in which the term is used

or how it is combined with other terms. Similar observations have been made by other researchers in the area of translating psychological assessment instruments into the Zulu language (Buntting & Wessels, 1991; John, 1996). The strategy adopted by the committee for dealing with such translation difficulties involved the provision of context and redundancy, as recommended by Brislin (1986). As a result, certain items are somewhat longer in the Zulu translation than in the original English.

With regard to the response categories, three of the five categories were identical across the two Zulu translations. The remaining two categories were discussed in the committee and agreement reached as to the most appropriate translations. The committee members were in agreement that the translated 5-point scale (a) captured the information contained in the original English, and (b) was appropriate for the target population. Alternative response formats, including two- and three-choice formats, were considered by the committee. The consensus view was that the original five-choice format would allow respondents to describe their symptoms with greater accuracy than would a simplified response format. The committee agreed that verbal administration of the SCL-90-R would require that the test administrator repeat the response categories for each item during the testing, since the capacity of respondents to retain the five categories in memory should not be assumed.

Upon completion of this step, a single Zulu version of the SCL-90-R had been produced. This version was used in the pretest study.

CHAPTER 5

THE PRETEST STUDY

The purpose of pretesting was to assess the comprehensibility and acceptability of the Zulu inventory in a small sample of people who were broadly representative of the general adult Zulu population, particularly in terms of educational level and English language proficiency. This was regarded as a means of compensating for the general reliance on bilingual translators and students, who were highly educated relative to the general population, in other phases of the study. This phase was also seen as an opportunity to identify any problems associated with verbal administration of the SCL-90-R.

5.1 METHOD

5.1.1 Subjects

The pretest study was conducted with workers at the University of Natal Research Farm, Ukulinga. The sample consisted of 12 subjects (9 males and 3 females) who volunteered to participate in the study. The mean age of the subjects was 38.5 years (SD = 12.9; range = 23 - 58). The mean number of years of formal schooling was 5 years (SD = 3.25; range: 0 - 10). Each of these subjects described themselves as having, at best, a rudimentary knowledge of the English language.

5.1.2 Procedures

Interviewing was conducted by two Zulu research assistants who were undergraduate psychology students. Prior to pretesting these assistants attended a training session with the author during which they were familiarised with the SCL-90-R and the

purpose of the pretest study. They were also briefed on procedures for interviewing and recording responses.

Informed consent was obtained verbally. Subjects were informed of the purpose of the research, the confidential manner in which their responses would be used, that the researchers were not in a position to offer any treatment or financial assistance, and that their decisions regarding whether or not to participate in the study would have no bearing on their subsequent employment prospects or treatment by their employers.

Each subject was interviewed individually and privately by a research assistant, the inventory being administered verbally. Due to the length of the SCL-90-R and constraints on the length of time that could be spent with each subject (1 hour), it was not practical to have each subject complete the entire inventory and also answer questions about each item. It was consequently decided to split the sample into three groups, four subjects responding to the first 30 items, four to the second 30, and four to the final 30 items. Each subject first responded to the 30 items and was then asked to go over the items again, explaining what each item meant and giving examples. These explanations were recorded by the research assistants and subsequently classified by the author, in collaboration with the assistants, according to whether or not they indicated that the item had been correctly understood. The research assistants were also instructed to record subjects' affective reactions to items (e.g., embarrassment, anger) and any other potentially relevant observations.

Items which were identified as presenting problems for the respondents were reviewed, and the necessary revisions made, by a committee consisting of the researcher and the three Zulu psychologists. Where necessary, revisions were also made to the English version of the SCL-90-R in order to clarify the meaning of the English items for the purpose of the bilingual technique.

5.2 RESULTS

Five items (numbers 5, 9, 12, 62, 78) were identified as possibly requiring revision in order to make the items more comprehensible or less offensive to respondents. The remaining items were found to be well understood and acceptable to the pretest sample. No problems relating to verbal administration of the inventory or to the response format were identified.

5.3 DISCUSSION

The five items were reviewed and dealt with by the committee as follows:

Item number 5, "Loss of sexual interest or pleasure" (*ukuphelelwa uthando noma ubumnandi ngokocansi*), a Depression scale item, tended to elicit some embarrassment or shyness on the part of the respondents. This had been anticipated by the committee during its assessment of the initial translations and back-translations, and considerable attention had been given to attempting to word the translation in such a way as to retain the meaning of the original English while minimising the potential to offend. While the committee agreed that this objective had been achieved, it was regarded as inevitable that the item would nevertheless cause some embarrassment with Zulu individuals, given that such matters are not usually discussed with strangers, particularly of the opposite sex. While not seeking to embarrass respondents, the significance of reduced sexual interest with respect to the diagnosis of a major depressive episode in other cultures (American Psychiatric Association, 1994, p.321) was regarded by the committee as providing justification for the inclusion of this item, which may prove to be of value in the clinical assessment of Zulu people. It should also be noted that the embarrassment is likely to have been intensified by the intrusive nature of the pretesting, respondents being required not only to say how much

they had been bothered by the problem but also to explain their responses to the interviewer. Such explanations would not be required in terms of the standard administration of the inventory. In view of the above considerations, the committee agreed that modification of this item was not required.

Item number 9, "Trouble remembering things" (*ukuba nenkinga yokukhumbula*) was frequently interpreted as referring to the experience of being troubled by memories, such as being preoccupied with painful memories of absent loved ones. This interpretation had not been foreseen by the translators or other committee members, and was clearly at odds with the intended meaning of the original item which could be paraphrased as "being forgetful". This item was revised by the committee to read *inkinga yokukhohlwa* which translates into the English, "a problem of forgetfulness". The ambiguity inherent in both the original English item and the Zulu translation was thus removed.

Item number 12, "Pains in heart or chest" (*izinhlungu enhliziyweni noma esifubeni*) was correctly interpreted by some subjects as referring to physical pain in the heart or chest. However, other subjects misinterpreted the item as referring to *emotional* pain, giving examples of being disappointed or having interpersonal problems. Since the phrase *izinhlungu enhliziyweni* (pains in the heart) is a popular Zulu idiom referring to emotional pain, the source of confusion was easily identified by the committee. Ironically, this was one of eleven items which had been identically translated in the two initial translation phases. This congruence, together with literal back-translations which matched the original English item, had apparently resulted in a false sense of confidence in the quality of the translation. The item had consequently been accepted with only minimal consideration by the committee. After reconsidering the item in the light of the pretesting, it was decided by the committee to delete the reference to the heart and thus simplify the item to read *izinhlungu esifubeni*

("Pains in the chest"). It was believed that this would reduce the likelihood of the item being misinterpreted, while the item was still likely to be endorsed by individuals experiencing (physical) pains in the area of the heart.

Item number 62, "Having thoughts that are not your own", an item on the Psychoticism subscale and suggestive of thought insertion, had been translated as *ukuba nemicabango engeyona eyakho*. Many of the respondents endorsed this item, having interpreted it as referring to being concerned about other people's problems or being sympathetic towards others. The psychiatric intent of the item was clearly being missed, with obvious implications for the validity of the item. The committee dealt with this problem by adding context and redundancy to the item in order to make the idea of thought insertion more explicit. The revised item read *ukuba nemicabango ekungesiyona eyakho, ekungathi ifakwe abanye abantu emqondweni wakho*, which translates into the English "Having thoughts that are not your own, as if they have been put into your mind by others".

Item number 78, "Feeling so restless you couldn't sit still", which had been translated as *ukuzizwa unexhala, uhluleka ngisho ukuhlala unganyakazi*, elicited a variety of explanations which indicated that the respondents were confused as to the meaning of the item. For example, one respondent explained the item as meaning: "if you are afraid, you remain seated to avoid what is frightening you". Since the literal translation was apparently misunderstood, the committee modified the item and added redundancy, in the form of a common Zulu idiom (literally translatable as "to have no buttocks") referring to restlessness/being unable to sit still. The revised item read *ukuzizwa ungaphumule, ungenasinqe, kangangoba ungakwazi nokuhlala uzothe* ("Feeling so restless, with no buttocks, that you couldn't sit still"). It was agreed by the committee that the revised item was culturally appropriate and likely to be understood by a wide range of Zulu people.

These examples demonstrate the importance of pretesting in identifying and guiding the revision of problematical items. Certain items which had apparently been well translated (based on back-translations which matched the original English, matching independent Zulu translations of the item, and agreement by bilingual experts that equivalence had been achieved) were nevertheless not well understood by members of the target population. These items required further revision. The recording of respondents' explanations of item meanings proved to be invaluable in identifying the sources of confusion and suggesting how the items should be revised. These findings are consistent with previous reports (Brislin et al., 1973; Kortmann, 1987; Kortmann & Ten Horn, 1988) supporting the use of pretesting procedures in addition to translation procedures such as back-translation. The pretesting also served to verify the translation quality and cultural appropriateness of the vast majority of the translated items, which were apparently well understood by the respondents.

The Zulu translation of the SCL-90-R is shown in Appendix B. The SCL-90-R is fully protected by copyright and may not be reproduced in any manner without the prior written permission of the author, L.R. Derogatis.

CHAPTER 6

PSYCHOMETRIC EVALUATION OF THE ZULU SCL-90-R

The purpose of this study was to investigate the reliability of the Zulu SCL-90-R, and its equivalence to the decentered English version. Metric equivalence has been achieved when the two versions of an instrument display comparable psychometric properties at both the item and scale level (Berry, 1980; Butcher & Pancheri, 1976; Clark, 1985). The methods used are presented, followed by the results and discussion.

6.1 METHODS

6.1.1 The Bilingual Retest Technique

6.1.1.1 *Subjects*

The subjects were Zulu first language students who were taking Zulu Mother-Tongue (ZMT) courses I and II at the Department of Zulu, University of Natal, Pietermaritzburg. The mean age of this sample was 23.8 years (SD = 3.5 years), and 59% were male. Completed inventories in both languages were received from 61 subjects.[2]

6.1.1.2 *Procedure*

The inventories were group administered during lecture time, with seven days elapsed between testings. In an effort to control for any order effects, the order of administration was reversed for the two groups, ZMT I (n = 40) responding to the

---

[2]Eighty-four Zulu and 87 English versions of the test were completed. However, for the purposes of the bilingual technique, only matched data from those 61 individuals who had been present at both testings were used.

Zulu version first and ZMT II (n = 21) taking the English version first. The researcher and two Zulu research assistants were present during the testing to answer any questions. Prior to testing, the instructions and the example were explained with the aid of an overhead projector. The purpose of the research was briefly explained to the subjects as relating to "the adaptation of a psychological test for use with Zulu people". The confidential nature of the study was also explained.

The data were analyzed using SPSS. In order facilitate statistical analysis and to make full use of the data collected, missing responses was replaced with the computed mean for the particular item. To retain missing data would have been problematic since, for subscale analyses on SPSS, all subjects with any missing data would have been excluded, substantially reducing the sample size. The missing responses comprised just 1.04% of the total number of responses, and were not concentrated on particular items or on either of the two language versions.

The data were analyzed as follows:

a) The internal consistencies (coefficient alpha) of each subscale were computed separately for the Zulu and English versions.
b) Item-subscale correlations were computed for each of the nine symptom dimensions of the Zulu and English versions.
c) Scores on the two versions, at both the scale and item level, were compared. This involved correlational analysis and consideration of mean score discrepancies. The effects of the order of administration and bilingual response sets were also investigated.

All raw data are available on request from the author.

6.1.2 Test-Retest Reliability

6.1.2.1 *Subjects*

The sample was comprised of Zulu students of the ZMT III class at the University of Natal, Pietermaritzburg. The eventual sample was rather small (N = 18), consisting of those individuals who were present for both administrations of the test. The mean age of the sample was 25.8 years (SD = 5.2), and 61% were male.

6.1.2.2 *Procedure*

The Zulu SCL-90-R was administered twice to the sample, with seven days elapsed between testings. In all other respects the procedures were the same as those followed for the bilingual technique.

As with the bilingual technique sample, missing data was replaced with the applicable item mean. Once again, missing responses constituted approximately one percent of the total responses.

The test-retest reliability coefficients and scale-level mean scores were computed for the Zulu version based on these data.

6.1.3 Factor Analysis

6.1.3.1 *Subjects*

Data from the ZMT I, II, and III groups were combined for the factor analyses. For each language version, data were analyzed for all subjects who had completed that version on their first testing (or only testing, if they were present for only one of the testings). This was done in order to maximise the size of the sample which completed the Zulu version, while avoiding any retest effects. This resulted in two separate

groups: 96 students who completed the Zulu version, and 45 who completed the English version.

6.1.3.2 *Procedure*

Responses to the two versions were subjected to separate factor analyses using SPSS. Principal-components analyses of the scale intercorrelations were conducted to assess whether the obtained factorial structure of the Zulu translation conformed to the hypothesised nine-dimensional structure (Derogatis, 1983), and to compare the structures obtained for the English and Zulu versions.

6.2 RESULTS

6.2.1 Internal Consistency

The internal consistency reliabilities (coefficient alpha) for the Zulu and English versions of the SCL-90-R, based on the responses of 61 bilingual students, are shown in Table 1.

Table 1.

Internal Consistency Coefficients

for the Zulu and English Versions of the SCL-90-R

Symptom Dimension	Number of items	Coefficient α	
		Zulu	Eng
Somatization	12	.77	.83
Obsessive-Compulsive	10	.80	.79
Interpersonal Sensitivity	9	.88	.78
Depression	13	.84	.83
Anxiety	10	.81	.83
Hostility	6	.83	.71
Phobic Anxiety	7	.63	.73
Paranoid Ideation	6	.74	.69
Psychoticism	10	.76	.80

6.2.2 Item-Subscale Correlations

Item-subscale correlations were computed individually for each of the nine symptom dimensions of the Zulu and English versions of the SCL-90-R, based on the sample of 61 bilingual students. These correlations were "corrected", each item being correlated with the sum of scores of all other items comprising the dimension. The ranges of corrected correlations for the items within each dimension are shown in Table 2. For both versions, each item correlated positively with the total of the symptom dimension to which it belongs.

Table 2.

Item-Subscale Correlations (Ranges)

for the Zulu and English Versions of the SCL-90-R

Symptom Dimension	Number of Items	Item-Subscale Correlations (Ranges)	
		Zulu	Eng
Somatization	12	.24-.63	.30-.64
Obsessive-Compulsive	10	.28-.75	.19-.63
Interpersonal Sensitivity	9	.39-.81	.26-.67
Depression	13	.17-.79	.16-.71
Anxiety	10	.20-.65	.40-.63
Hostility	6	.57-.66	.37-.55
Phobic Anxiety	7	.21-.61	.31-.57
Paranoid Ideation	6	.28-.62	.36-.51
Psychoticism	10	.11-.65	.25-.58

6.2.3 Test-Retest Reliability

Test-retest reliability coefficients for the Zulu SCL-90-R are shown in Table 3. These results are based on a sample of normal bilingual students (N = 18) who completed the Zulu version twice, with seven days elapsed between testings. The coefficients range from .46 (Paranoid Ideation) to .90 (Phobic Anxiety).

Table 3.

Test-Retest Reliability Coefficients for the Zulu SCL-90-R

Symptom Dimension	Test-Retest Reliability
Somatization	.72
Obsessive-Compulsive	.84
Interpersonal Sensitivity	.86
Depression	.79
Anxiety	.73
Hostility	.75
Phobic Anxiety	.90
Paranoid Ideation	.46
Psychoticism	.58

6.2.4 Factor Analysis

Scree plots of the English and Zulu factor analyses are shown in figures 1 and 2 respectively. For both versions only one factor emerged with an eigenvalue of greater than one. For the English version, this first factor accounted for 71.2% of the total variance. The next largest factor accounted for just 6.4% of the variance. For the Zulu version, the first factor accounted for 64.7% of the variance, and the second factor for 7.1%.

Factor loadings of the nine symptom dimensions on the first factor, for both the English and Zulu versions, are shown in Table 4. For both versions, all nine scales loaded strongly on the first factor. These correlations range from .77 (Hostility) to .91 (Anxiety) for the English version, and .72 (Somatization) to .90 (Anxiety) for the Zulu version.

Table 4.

Factor Loadings of English and Zulu

SCL-90-R Symptom Dimensions on Factor 1

	Loading on Factor 1	
Symptom Dimension	English (n = 45)	Zulu (n = 96)
Somatization	.81	.72
Obsessive-Compulsive	.89	.83
Interpersonal Sensitivity	.81	.83
Depression	.88	.87
Anxiety	.91	.90
Hostility	.77	.77
Phobic Anxiety	.82	.73
Paranoid Ideation	.84	.81
Psychoticism	.85	.84

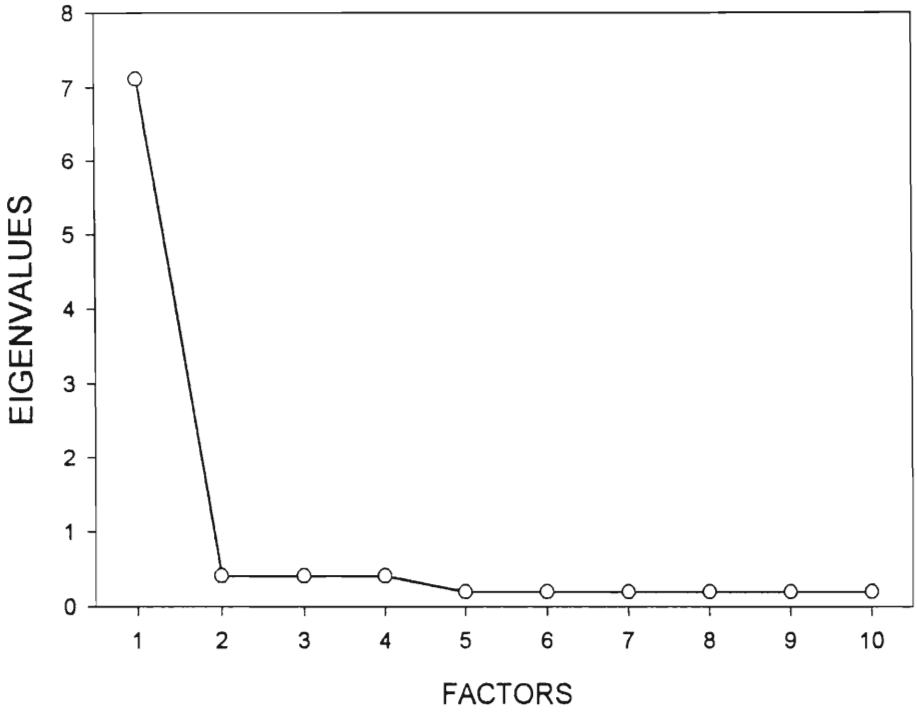

Figure 1 Eigenvalues of factors extracted for English SCL-90-R

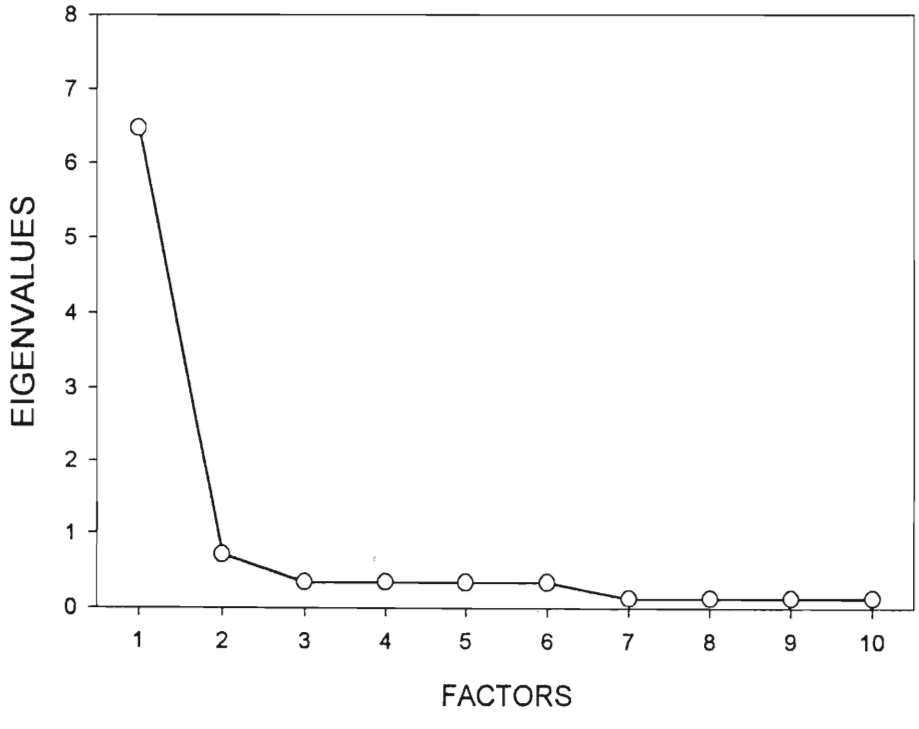

Figure 2 Eigenvalues of factors extracted for Zulu SCL-90-R

6.2.5 Order Effects and Response Sets

The nine symptom dimensions and Global Severity Index (GSI) scores were analyzed in order to assess the effects of the order of administration and other possible response sets on the results. This was done in three ways. Firstly, mean scores on the first and second administration were compared for the Zulu-Zulu test-retest sample (N = 18). This allowed for a relatively pure assessment of the effect of the order of administration on the Zulu version scores, without the influence of effects due to different languages being administered. These results, for the nine symptom dimensions and the GSI, are shown in Table 5.

Secondly, the total bilingual administration sample of 61 subjects was analyzed according to whether their scores differed between the first and second administrations, regardless of which language was administered first. These results are shown in Table 6.

Finally, the scores were analyzed separately for the two bilingual administration groups, considering both the order and language of administration. These results are shown in Table 7. Group A (n = 40) received the Zulu translation first, while Group B (n = 21) received the English version first.

Figure 3 illustrates the comparative GSI mean scores on the English and Zulu versions, for Group A and Group B. The GSI represents the mean score of all items on the SCL-90-R and is thus a good summary measure.

Table 5.

Scale Means and Discrepancies

by Order of Administration

for the Zulu-Zulu Test-Retest Sample

Scale	Administration		Discrepancy (1st - 2nd)	Significance[a]
	1st	2nd		
SOM	.95	.75	+.20	NS
O-C	1.32	.92	+.40	**
INT	1.27	1.05	+.22	NS
DEP	1.14	1.00	+.14	NS
ANX	.87	.77	+.10	NS
HOS	1.11	1.00	+.11	NS
PHOB	.75	.63	+.12	NS
PAR	.99	.92	+.07	NS
PSY	.87	.78	+.09	NS
GSI	1.05	.87	+.13	*

SOM = Somatization, O-C = Obsessive-Compulsive, INT = Interpersonal Sensitivity, DEP = Depression, ANX = Anxiety, HOS = Hostility, PHOB = Phobic Anxiety, PAR = Paranoid Ideation, PSY = Psychoticism, GSI = Global Severity Index

[a]2-tailed Wilcoxon matched-pairs signed-ranks test

*p<.05 **p<.01

Table 6.

Scale Means and Discrepancies

by order of Administration

for the Combined Bilingual Administration Sample

Scale	Administration		Discrepancy (1st - 2nd)	Significance[a]
	1st	2nd		
SOM	.79	.76	+.03	NS
O-C	1.34	1.28	+.06	NS
INT	1.37	1.26	+.11	NS
DEP	1.27	1.20	+.07	NS
ANX	.84	.87	-.03	NS
HOS	1.08	.93	+.15	NS
PHOB	.66	.65	+.01	NS
PAR	1.12	1.07	+.05	NS
PSY	.91	.89	+.02	NS
GSI	1.05	.99	+.06	NS

[a]2-tailed Wilcoxon matched-pairs signed-ranks test

Table 7.

Scale Means and Discrepancies by Order of Administration for Groups A (Zulu first) and B (English first)

Group A: Zulu First (n = 40)				
	Administration		Discrepancy	
Scale	1st (Zulu)	2nd (Eng)	(Zul - Eng)	Significance[a]
SOM	.83	.88	-.05	NS
O-C	1.34	1.42	-.08	NS
INT	1.46	1.42	+.04	NS
DEP	1.30	1.35	-.05	NS
ANX	.92	1.01	-.09	NS
HOS	1.16	1.04	+.12	NS
PHOB	.70	.81	-.11	NS
PAR	1.19	1.31	-.12	NS
PSY	.97	1.10	-.13	NS
GSI	1.10	1.14	-.04	NS
Group B: English First (n = 21)				
	Administration		Discrepancy	
Scale	1st (Eng)	2nd (Zulu)	(Eng - Zul)	Significance[a]
SOM	.71	.52	+.19	*
O-C	1.35	1.01	+.34	**
INT	1.19	.96	+.23	NS
DEP	1.23	.91	+.32	**
ANX	.69	.61	+.08	NS
HOS	.91	.81	+.10	NS
PHOB	.58	.36	+.22	*
PAR	.99	.61	+.38	***
PSY	.79	.49	+.30	***
GSI	.95	.70	+.25	***

[a]2-tailed Wilcoxon matched-pairs signed-ranks test

*p<.05 **p<.01 ***p<.001

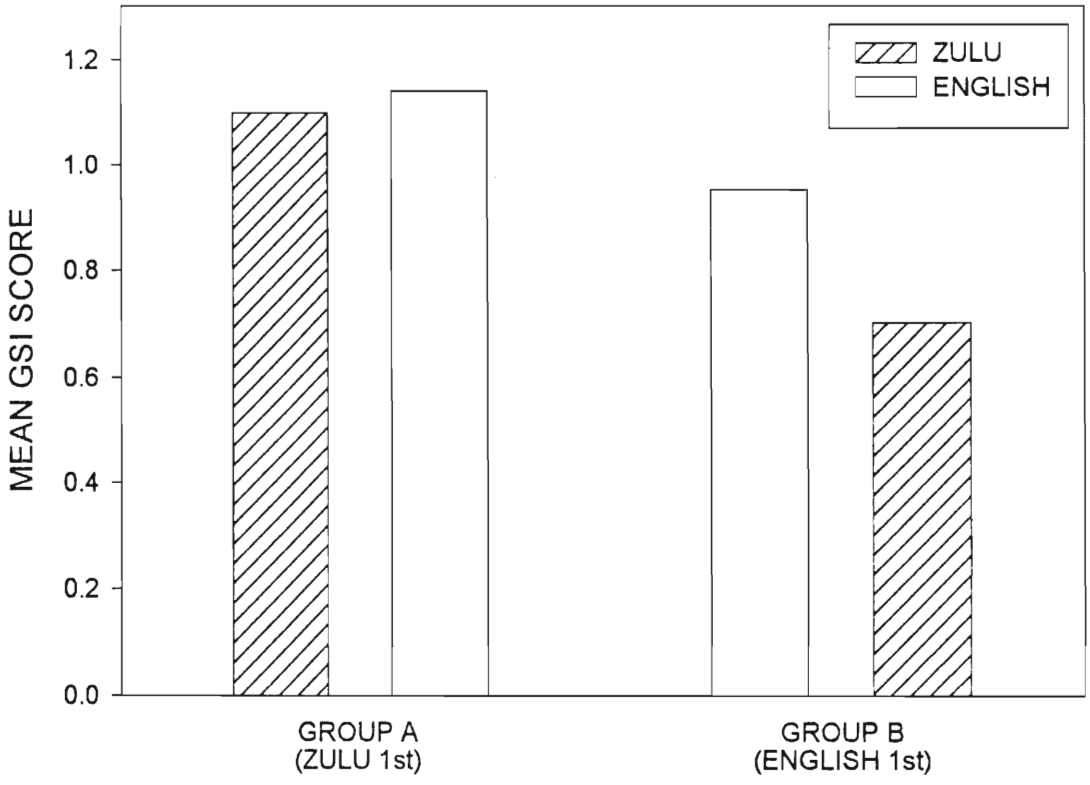

Figure 3 Comparison of mean GSI scores for Group A and Group B

6.2.6 The Bilingual Technique

The scale-level bilingual technique results (N = 61) are shown in Table 8. Scale-level equivalence was assessed in terms of the correlational relationship and mean score discrepancies between the two versions.

Table 8.

Scale Means, Discrepancies, and Correlations for the Zulu and English Versions of the SCL-90-R

	Scale Mean				
Scale	Zulu	Eng	Discrepancy (Zul - Eng)	Significance[a]	r[b]
SOM	.72	.82	-.10	NS	.70
O-C	1.23	1.40	-.17	*	.69
INT	1.29	1.34	-.05	NS	.82
DEP	1.16	1.31	-.15	*	.76
ANX	.81	.90	-.09	NS	.71
HOS	1.00	.99	+.01	NS	.67
PHOB	.58	.73	-.15	NS	.65
PAR	.99	1.20	-.21	**	.63
PSY	.80	.99	-.19	**	.75
GSI	.97	1.07	-.10	*	.80

[a]2-tailed Wilcoxon matched-pairs signed-ranks test

*p<.05 **p<.01

[b]p<.001 for all correlations

The item-level bilingual test-retest results (N = 61) are shown in Table 9. Equivalence between the English and Zulu items was assessed both in terms of item-item correlations and discrepancies between the mean scores across the two versions.

Table 9.

Item Means, Discrepancies, and Item-item Correlations for the 90 Zulu and English Items of the SCL-90-R

Item	Item Mean		Item-item Discrepancy (Zul - Eng)	Significance[a]	Item-item Correlation
	Zulu	Eng			
1	1.44	1.49	-.05	NS	.62
2	1.11	.97	+.14	NS	.39
3	1.33	1.36	-.03	NS	.29
4	.65	.57	+.08	NS	.43
5	1.11	1.13	-.02	NS	.40
6	.69	1.34	-.65	***	.14
7	1.20	.80	+.40	*	.36
8	.75	.98	-.23	NS	.26
9	1.73	1.67	+.06	NS	.61
10	.85	1.66	-.81	***	.28
11	1.85	1.76	+.09	NS	.46
12	.45	.75	-.30	NS	.31
13	.43	.61	-.18	NS	.26
14	1.48	1.52	-.04	NS	.44
15	.15	.46	-.31	**	.55
16	.18	.30	-.12	NS	.52
17	.28	.58	-.30	*	.00
18	.80	1.38	-.58	**	.24
19	.93	.90	+.03	NS	.05
20	.74	.98	-.24	NS	.67
21	1.33	1.10	+.23	NS	.42
22	1.25	.79	+.46	**	.33
23	.55	.60	-.05	NS	.50
24	1.56	.92	+.64	**	.39
25	.31	.64	-.33	*	.27
26	1.44	1.44	0.00	NS	.42
27	.69	.88	-.19	NS	.77

Table 9 (continued)

Item Means, Discrepancies, and Item-item Correlations for the 90 Zulu and English Items of the SCL-90-R

Item	Item Mean		Item-item Discrepancy (Zul - Eng)	Significance[a]	Item-item Correlation
	Zulu	Eng			
28	1.21	1.25	-.04	NS	.35
29	1.52	1.52	0.00	NS	.59
30	1.39	1.68	-.29	NS	.53
31	1.74	2.00	-.26	NS	.48
32	1.18	1.47	-.29	NS	.47
33	.52	.77	-.25	NS	.25
34	1.82	1.38	+.44	*	.40
35	.72	1.02	-.30	NS	.40
36	1.41	1.36	+.05	NS	.32
37	1.27	1.34	-.07	NS	.72
38	1.51	1.90	-.39	NS	.26
39	.93	1.05	-.12	NS	.27
40	1.05	.97	+.08	NS	.23
41	1.28	1.43	-.15	NS	.27
42	1.18	.76	+.42	**	.46
43	1.23	.97	+.26	NS	.44
44	.93	1.20	-.27	NS	.17
45	1.33	1.30	+.03	NS	.26
46	1.56	1.53	+.03	NS	.52
47	.70	.81	-.11	NS	.43
48	.34	.38	-.04	NS	.44
49	.66	.60	-.06	NS	.13
50	1.07	.89	+.18	NS	.16
51	.60	1.07	-.47	**	.22
52	.59	.69	-.10	NS	.23
53	.52	.70	-.18	NS	.62
54	1.05	1.40	-.35	*	.62

Table 9 (continued)

Item Means, Discrepancies, and Item-item Correlations for the 90 Zulu and English Items of the SCL-90-R

Item	Item Mean		Item-item Discrepancy (Zul - Eng)	Significance[a]	Item-item Correlation
	Zulu	Eng			
55	1.10	1.36	-.26	NS	.37
56	.69	1.15	-.46	**	.21
57	1.72	.98	+.74	***	.42
58	.41	.92	-.51	***	.64
59	.61	.97	-.36	*	.42
60	1.07	.75	+.32	NS	.52
61	1.75	1.46	+.29	NS	.68
62	.56	.78	-.22	*	.49
63	.70	.66	+.04	NS	.68
64	1.26	.98	+.28	NS	.58
65	1.05	.87	+.18	NS	.20
66	1.13	.86	+.27	NS	.26
67	.39	.64	-.25	*	.61
68	1.15	1.34	-.19	NS	.55
69	1.18	1.24	-.06	NS	.40
70	.80	1.00	-.20	NS	.56
71	1.26	1.48	-.22	NS	.02
72	.63	1.08	-.45	*	.21
73	.85	1.42	-.57	**	.40
74	.74	1.34	-.60	**	.25
75	.38	.61	-.23	NS	.16
76	1.18	1.28	-.10	NS	.44
77	.83	.95	-.12	NS	.58
78	.64	.82	-.18	NS	.27
79	.80	1.09	-.29	NS	.54
80	1.13	1.23	-.10	NS	.38
81	.77	.64	+.13	NS	.15

Table 9 (continued)

Item Means, Discrepancies, and Item-item Correlations for the 90 Zulu and English Items of the SCL-90-R

Item	Item Mean		Item-item Discrepancy (Zul - Eng)	Significance[a]	Item-item Correlation
	Zulu	Eng			
82	.38	.57	-.19	NS	.48
83	.80	1.23	-.43	*	.27
84	1.13	1.14	-.01	NS	.67
85	.97	1.38	-.41	**	.61
86	.61	.90	-.29	NS	.28
87	.98	1.33	-.35	NS	.24
88	.75	1.23	-.48	*	.33
89	1.10	1.02	+.08	NS	.29
90	.69	1.00	-.31	NS	.47

[a]2-tailed Wilcoxon matched-pairs signed-ranks test

*$p<.05$ **$p<.01$ ***$p<.001$

6.3 DISCUSSION

6.3.1 Internal Consistency

Coefficient alpha estimates the reliability of a scale based on the average correlation among items within the scale. It is a measure of the consistency with which the items selected to represent a particular scale actually reflect the same underlying factor or symptom construct. According to Nunnally (1978), coefficient alpha provides a good estimate of a scale's reliability in most situations and should be applied to all new measurement methods. In addition, in the context of the current study, comparison of the coefficients obtained for the two language versions provides an index of the effect of the translation on the instrument's reliability.

A reliability coefficient of at least .75 is frequently recommended for psychological scales (Shaw & Wright, 1967). The coefficients originally reported for the SCL-90-R, based on 219 "symptomatic volunteers" in the United States (Derogatis, 1983), range from .77 (Psychoticism) to .90 (Depression). The coefficients obtained for the Zulu and English versions in the current study (see Table 1, p. 68), while somewhat lower than those reported by Derogatis (1983), are generally satisfactory. For the Zulu version, seven of the nine scales satisfy the minimum criterion of .75, while for the English version six of the scales are satisfactory.

Those dimensions with satisfactory coefficients alpha in the one language tend to have similarly satisfactory measures in the other, while the Phobic Anxiety and Paranoid Ideation dimensions show relatively poor internal reliability across both languages. The Hostility dimension is the exception, the coefficient for the Zulu version (.83) being considerably higher than that for the English version (.71). The average coefficient for the nine scales is equivalent for the English and Zulu versions, working out to .78. Overall, these results

suggest that the translation of the SCL-90-R into Zulu had little effect on its internal reliability.

It is noteworthy that the Phobic Anxiety, Paranoid Ideation, and Hostility scales are the shortest of the SCL-90-R, consisting of six or seven items each. Since the size of the reliability coefficient is affected by the number of items (Nunnally, 1978), the low coefficients alpha may be attributable, in part, to the relatively low number of items comprising these scales. In addition, consideration of the items comprising the Phobic Anxiety and Paranoid Ideation scales suggests possible reasons why, as applied to the target population, these scales may not consistently measure a single underlying dimension.

According to Derogatis (1983), the Phobic Anxiety dimension is intended to "focus on the more pathognomic and disruptive manifestations of phobic behaviour. The actual structure of the dimension is in close agreement with the definition of agoraphobia" (p. 9). However, in the context of ongoing violence in KwaZulu-Natal, certain items (e.g., No. 25: "feeling afraid to go out of your house alone"; No. 47: "feeling afraid to travel on buses or trains") may not be reliable indices of phobic behaviour. As previously discussed, such fears may be regarded as normal or adaptive responses to somewhat abnormal circumstances. However, other items (e.g., No. 82: "feeling afraid you will faint in public") may more reliably measure what the scale is designed to assess. Consequently, the relatively poor internal consistency of the Phobic Anxiety dimension may reflect that, due to pervasive situational variables affecting the target population, the items contained in this subscale may not be measuring a single symptom construct.

Analysis of the Paranoid Ideation dimension similarly suggests reasons why, as applied to Zulu people, this scale may produce poor measures of internal consistency. As discussed by Ngubane (1977) and Buntting and Wessels (1991), traditional

African beliefs regarding causality maintain that there is typically an agent of harm, one's illness or misfortune being attributable to the malevolent intentions of others. A study by Edwards, Grobbelaar, Nene, Makunga, Kunene, and Sibaya (1985) reported that such traditional beliefs were held by the majority of both rural and urban Zulus, including university students. Items such as No. 8: "feeling others are to blame for most of your troubles" may consequently be of questionable validity as a measure of paranoid thinking among Zulu people, measuring instead the degree to which traditional African beliefs are held by the respondents. By comparison, another item on the Paranoid Ideation scale, No. 18: "feeling that *most* people cannot be trusted" (emphasis added), may be a more valid measure of paranoid thinking, a normal cultural belief being distinguishable from delusional paranoia according to the number of people who are perceived by the individual as having harmful intentions (Buntting and Wessels, 1991). It thus seems plausible that strong cultural factors may explain the relatively poor internal consistency of the Paranoid Ideation scale.

The above discussion emphasises the need to consider contextual and cultural influences in the assessment process, and the importance of establishing the criterion validity of the SCL-90-R as applied to Zulu people.

The current study was based on a nonclinical population whereas the original internal consistency coefficients (Derogatis, 1983) were based on "symptomatic volunteers". Since the SCL-90-R was designed primarily to reflect the psychological symptom patterns of psychiatric and medical patients (Derogatis, 1983), it seems plausible that the scales may have lower internal consistency as applied to a nonclinical population by virtue of lower levels and different patterns of symptomatology in the group. The legitimacy of a simple comparison of the results of the current study with those reported by Derogatis (1983) may thus be questioned. It is recommended that future

studies investigate the internal consistency of the Zulu SCL-90-R with psychiatric samples.

6.3.2 Item-Subscale Correlations

The higher an item correlates with a subscale score the stronger its relationship to the common factor among the items, and the more it adds to the reliability of the subscale (Nunnally, 1978). An item that correlates near zero or negatively with a subscale score is likely to be a poor measure of the construct that the subscale is measuring. It is possible that the item has been poorly translated, is culturally inappropriate, or is otherwise confusing or irrelevant to the respondents.

The correlations obtained for both versions in the current study (see ranges in Table 2, p. 69) are all positive and appear to be of acceptable levels. The absence of negative or zero item-subscale correlations for the Zulu version suggests that, in the sample employed, all items are relevant and comprehensible, and contribute substantially to the common variance among the items comprising the respective scales. These results are consistent with the generally satisfactory internal consistency measures shown in Table 1 (p. 68).

6.3.3 Test-Retest Reliability

The test-retest reliabilities of the Zulu SCL-90-R scales were investigated for two primary reasons. Firstly, the widespread use of the SCL-90-R for repeated assessments across time is supported by the high test-retest reliabilities reported for the original SCL-90-R (Derogatis, 1983). Similarly high test-retest reliabilities for the Zulu translation would support similar use of this version. Secondly, the Zulu-Zulu test-

retest coefficients may provide some basis for comparison when assessing the results of the bilingual test-retest study.

The test-retest reliabilities (see Table 3, p. 70) are generally of acceptable levels. The coefficients for the first seven scales of the Zulu SCL-90-R are comparable with those reported by Derogatis (1983), which ranged from .78 for Hostility to .90 for Phobic Anxiety. Lower correlations were obtained for the Psychoticism and Paranoid Ideation scales. The poor showing of the Paranoid Ideation scale (r_{tt} = .46) is consistent with the relatively poor result obtained on the internal consistency measure, and suggests that this scale may be of generally poor reliability as applied to Zulu people.

Bearing in mind the small number of cases and the inherent instability of psychological states and symptoms, the test-retest reliabilities obtained for the Zulu SCL-90-R are generally satisfactory. These findings await confirmation in future studies with larger samples and utilizing different groups, including psychiatric patients.

6.3.4 Factor Analysis

Factor analysis was conducted for two reasons. Firstly, in order to allow comparison of the obtained factor structure of the Zulu SCL-90-R with the hypothetical dimensional structure of the SCL-90-R (Derogatis, 1983) and the factor structures obtained in other studies. Secondly, to allow comparison of the factor structures obtained for the English and Zulu versions in the current study. Different factor structures across the two language versions would be suggestive of a qualitative change having occurred in the translation-adaptation process (Ben-Porath, 1990), raising questions concerning the equivalence of the translation and the cross-cultural validity of the SCL-90-R.

Comparison of the factor analysis results of the Zulu and English versions of the SCL-90-R reveal broadly similar factor structures, and no evidence to suggest that a qualitative change has resulted from the translation into Zulu. As shown in Figures 1 and 2 (p. 72), a single large factor emerged which accounted for most of the variance of both versions. Subsequent factors were much smaller than the first factor and had eigenvalues of less than 1. For both versions, all nine subscales loaded strongly on this first factor (see Table 4, p. 71).

Since principal components analysis extracted only one factor with an eigenvalue of greater than one, factor rotation was not performed.

These findings do not support the hypothesised multidimensional structure of the SCL-90-R (Derogatis, 1983). Instead, the results are consistent with those of previous studies (Carpenter and Hittner, 1995; Hoffmann and Overall, 1978; Holcomb et al., 1983; Schwartzwald et al., 1991) which suggest that the SCL-90-R is best utilised as a unidimensional measure of global psychological distress, and that any interpretation of subscale scores should be conducted with circumspection. This advice is accordingly extended to prospective users of the Zulu SCL-90-R.

The general failure of psychometric studies of the SCL-90-R to produce evidence of distinct multiple dimensions is perhaps not surprising when one considers the limitations of the categorical approach to the classification of mental disorders. The DSM-IV (APA, 1994), acknowledges these limitations, stating that "there is no assumption that each category of mental disorder is a completely discrete entity with absolute boundaries dividing it from other mental disorders" (p. xxii). Further, a given symptom is seldom specific to a particular disorder, as is empirically demonstrated in the item overlap

between scales on criterion-keyed assessment instruments such as the MMPI.

This factor analytic study was conducted at scale level and on a small number of nonpatient students. Due to low subject-to-variable ratios, item-level factor analysis was not attempted. As previously discussed, most researchers recommend a subject-variable ratio of at least 5 to 1 (Comrey, 1988). Applied to the SCL-90-R, this recommendation would necessitate samples of at least 450 subjects for item-level analysis. Such sample sizes were beyond the resources of the current study. Also, since Carpenter and Hittner (1995) reported gender differences in the number and composition of SCL-90-R symptom dimensions, the legitimacy of factor-analyzing the responses of a combined group of men and women is questionable. It is recommended that further investigations be conducted into the factor structure of the Zulu SCL-90-R. Such investigations should involve different populations, including men and women and psychiatric inpatients and outpatients, and should ideally use large enough samples to permit item-level analysis.

6.3.5 Order Effects and Response Sets

In employing the bilingual technique (Prince & Mombour, 1967), it would seem essential that the data be assessed for any effects due to the order of administration or bilingual response sets. If such effects exist, failure to take account of them may result in score discrepancies between the two versions, which may be misinterpreted as being due to translation errors. In the current study, in an attempt to control for order effects, one group of subjects (Group A; n = 40) received the Zulu version first and another group (Group B; n = 21) received the decentered English version first. Ideally, the two groups should have been more similar in size, the difference being due to different student numbers in the first- and second-year courses.

In order to investigate the presence and extent of effects due to the order of administration, the scale scores of the Zulu-Zulu test-retest sample (N = 18) were analyzed for differences between the first and second administrations. These results are shown in Table 5 (p. 74). All nine symptom dimensions scored lower on the second administration, the discrepancy being statistically significant for the Obsessive-Compulsive dimension. The discrepancy between the GSI scores on first and second administrations indicates an overall order effect, with lowered scores on the second administration. This effect is consistent with the findings of John (1996) in his study of a Zulu translation of the General Health Questionnaire. Such a retest effect has implications for the use of the Zulu SCL-90-R for repeated assessments over time, such as in the monitoring of change due to therapeutic intervention. Failure to take account of this effect may tend to suggest therapeutic improvement where none exists, or may exaggerate the extent of any actual improvement. Consequently, if the Zulu SCL-90-R is to be used for such purposes, it is suggested that norms be established for repeated administrations.

With regard to the original SCL-90-R, Derogatis (1983) states "we have been unable to detect any significant 'practice' effects which might bias the profile on repeated administration" (p. 5). Unfortunately, while Derogatis (1983) reports satisfactory test-retest reliability correlations, he does not provide an analysis of discrepancies between mean scale scores. The current study suggests that satisfactory test-retest reliability coefficients are not sufficient to rule out significant retest effects. To the extent that a retest effect consistently affects all scales of the test, scale correlations may be high despite significant scale score discrepancies. This may be illustrated with reference to the Obsessive-Compulsive scale. This scale yielded a satisfactory test-retest correlation of .84 (see Table 2, p. 69) but showed a significant discrepancy ($p < .01$) between the first and second administrations (see Table 5, p. 74).

The order effect was further investigated by analyzing the results of the bilingual test-retest sample. First, the scale scores for the 61 subjects were analyzed for differences between the first and second administrations, irrespective of which language was administered first. These results are shown in Table 6 (p. 75). While the discrepancies between the mean scale scores were not statistically significant, all scales, with the exception of the Anxiety dimension, scored lower on the second administration. These results thus tend to confirm the retest effect found in the Zulu-Zulu test-retest sample.

The scale scores were subsequently analyzed separately for Groups A and B. These results are shown in Table 7 (p. 76). For Group A (Zulu first), while none of the discrepancies were statistically significant, eight of the ten scale scores were higher for the second (English) administration. The exceptions were the Interpersonal Sensitivity and Hostility dimensions, which were slightly higher for the first (Zulu) administration.

For Group B (English first), all scales were higher on the first (English) administration. These discrepancies were statistically significant for all except the Interpersonal Sensitivity, Anxiety, and Hostility dimensions.

With reference to Figure 3 (p. 77), which illustrates the comparative Global Severity Index mean scores on the English and Zulu versions, these results may be explained in terms of both an order effect and a language effect. With regard to the order effect, it has been established that respondents scored lower on the second administration. In terms of the language effect, it appears that the respondents scored higher when tested in English than when tested in Zulu.

Consequently, for Group A (Zulu first), the language effect (English higher) was apparently attenuated by the order effect (second administration lower). The net result was higher scores

on the English version, but of a magnitude which was not statistically significant.

However, for Group B (English first), the language effect (English higher) was augmented by the order effect (first administration higher), resulting in significantly higher scores on the English version.

Discrepancies between the scores obtained on the two language versions are consistent with previous studies which found that bilinguals' responses to items differed depending on the language employed (Bond & Yang, 1982; Ervin, 1964; Marin et al., 1983; Tyson et al., 1988; Yang & Bond, 1980). However, the direction of the difference in the current study is contrary to that reported by John (1996), who found that Zulu bilinguals scored lower on the English than the Zulu version of the GHQ. John (1996) offers the social desirability hypothesis (Marin et al., 1983) as an explanation for his results, respondents scoring lower (more "normal") in their second language in order to present themselves in a more socially desirable way. With regard to the results of the current study, a variety of possible explanations should be considered.

a) Bilingual Proficiency

One possible explanation is that the respondents were not fully proficient in their second language and may consequently have experienced difficulty in understanding the English version of the SCL-90-R[3]. Proficiency in the English language was not assessed, but was assumed on the basis that all subjects were students at an English-medium university. Also, while the Zulu translation was pretested to ensure its comprehensibility, the comprehensibility of the English version as applied to Zulu first-language respondents was not assessed. It thus seems

[3]The ideal procedure when using the bilingual technique may be to employ first-language speakers of each of the two languages. Since relatively few first-language English speakers are proficient in Zulu, this ideal was not feasible in the current study.

plausible that the sample may have experienced more difficulty in understanding the English version than the Zulu translation, resulting in discrepant responses.

If difficulty understanding the English version contributed to the language effect, it remains to be explained why subjects should have tended to report higher levels of distress in response to items which they did not understand. Since in African culture it is considered polite to answer in the affirmative if one does not understand a question, there may be some cultural basis for such a response set. Gillis et al. (1982) reported that Xhosa subjects tended to respond in the affirmative when they did not understand questions asked in the Present State Examination (PSE). The possibility that such an acquiescent response set may also occur with self-report inventories may help to explain the results of the current study and should be subjected to further investigation. A related possibility is that difficulty, or unfamiliarity, with the English language may have resulted in higher levels of test anxiety and thus tended to elicit higher scores in response to the English version.

Bilingual proficiency may also explain the lower scores for Group B relative to Group A on the English version of the test. Since the order and language effects tend to cancel each other for Group A but augment one another for Group B, one might expect that the mean GSI score on the English administration for Group B would be higher than that on either administration for Group A. The fact that this is not the case (see Figure 3, p. 77) may be due to greater proficiency in English for Group B. Due to logistical considerations, the inventories having been administered during lecture time, subjects were not randomised with regard to year of study. Group A consisted of students in a first year course (ZMT I) whereas Group B consisted of students in a second year course (ZMT II). On average, Group B subjects, as more senior students, may be expected to be more proficient in English than the students in Group A, the former

having had more exposure to the English language at university level. In addition, first year students who are particularly weak in English may be less likely to progress to second year level than their classmates who are more proficient in English. Consequently, it is plausible that Group B may have encountered less difficulty in understanding the English version of the SCL-90-R, resulting in a less powerful language effect and thus lower scores.

b) Social Desirability

As applied to bilingual responses, the social desirability hypothesis (Marin et al., 1983) maintains that bilinguals may tend to respond in a more socially desirable direction when answering in a second language. Tyson et al. (1988) have suggested that this social desirability effect may be more likely to occur when the culture associated with the second language is perceived as being of higher status than the respondent's own culture. It may be argued that, given South Africa's history of racial oppression and a university environment where the language of instruction is English and most lecturers are white and English speaking, respondents in the current study may perceive the English language-culture as being dominant. Accordingly, one may expect Zulu respondents to score in a more socially desirable manner when responding to the English version of the SCL-90-R. However, notwithstanding the argument put forward by John (1996) that lower scores on a psychiatric inventory are more "normal" and thus socially desirable, the symptom checklist format of the SCL-90-R may be less susceptible to such effects than would items dealing explicitly with attitudes or values. This view is supported by the higher scores on most scales when the English version of the SCL-90-R was administered, the opposite of what might be expected in terms of the social desirability hypothesis. The exceptions to this pattern are the Hostility and Interpersonal Sensitivity dimensions, which scored lower on the English version for Group A, and produced non-significant discrepancies for Group B (see Table 7, p. 76). Examination of the content of

these scales suggests that they may be more susceptible to social desirability response sets than the other scales. This is particularly evident in the case of the Hostility dimension, the items of which reflect urges or behaviours which may be universally regarded as socially undesirable. These include item number 24: "Temper outbursts you cannot control" and item number 63: "Having urges to beat, injure, or harm someone". The Interpersonal Sensitivity dimension contains items relating to comparing oneself to others and viewing others in a negative way. Thus, while most of the SCL-90-R dimensions tend to relate to somatic symptoms or internalised psychological distress, the Hostility and Interpersonal Sensitivity dimensions are more directly interpersonal in focus and may consequently be more sensitive to the social desirability response set.

The author was unable to find any studies relating to the social desirability of the items comprising the SCL-90-R. Future research focusing on this issue (e.g., by having Zulu subjects rate the desirability of the Zulu items) could serve to substantiate or negate the above argument.

c) Ethnic Affirmation and Cross-Cultural Accommodation

In terms of the ethnic affirmation hypothesis (Bond & Yang, 1980), being tested in a language other than one's first language may increase awareness of one's ethnicity and tend to elicit responses which affirm that ethnicity. Cross-cultural accommodation involves the opposite bias, subjects responding in the way which they perceive to be appropriate to the culture associated with the language of testing. These hypotheses may be expected to be of limited relevance in the case of the SCL-90-R items, which generally do not appear to deal explicitly with attitudes or values which are clearly more appropriate or affirming with regard to one culture than the other. Possible exceptions to this are the Psychoticism and Paranoid Ideation scales, both of which do contain items relating to beliefs which differ markedly across the two cultures concerned. The Psychoticism dimension contains items such as number 16:

"Hearing voices that other people do not hear", and number 35: "Other people being aware of your private thoughts". While such experiences or beliefs may be regarded as abnormal or pathological in terms of Western culture, they may be culturally validated by traditional African beliefs regarding the ability of the ancestors to communicate with the living, and the powers of certain people, such as *abangoma*, to know the thoughts of others. Similarly, as previously discussed, the Paranoid Ideation dimension contains items which may relate to widely-held traditional African beliefs regarding causality but which would be regarded as abnormal in Western culture.

In view of the above discussion, if response sets due to ethnic affirmation or cross-cultural accommodation played a role in the results of this study, it seems reasonable to expect that their influence would be strongest on the Paranoid Ideation and Psychoticism dimensions. Consideration of the mean scale scores for these two dimensions (see Table 7, p. 76) shows that both Group A and Group B scored higher on the English version, suggesting that ethnic affirmation may have occurred. For Group A any such effect may have been attenuated by the order effect (English having been administered second to these subjects), resulting in a non-significant discrepancy. For group B, the order effect would have been augmented by an ethnic affirmation effect, which may explain the particularly high discrepancies for the Paranoid Ideation and Psychoticism scales ($p < .001$).

Examination of the results reveals little evidence to suggest that cross-cultural accommodation played a substantial role in this study.

In summary, discrepancies between scores on the Zulu and English versions of the SCL-90-R may be plausibly explained in terms of an order effect, whereby subjects tended to score lower on their second testing, and a language effect, whereby subjects tended to score higher in English. The language effect may be due to the subjects, who were all first-language Zulus,

experiencing difficulty understanding the English version. This difficulty may have resulted in higher scores due to a culturally-related acquiescent response set. In addition, social psychological elements relating to social desirability and ethnic affirmation may have influenced responses to certain items, particularly on the Interpersonal Sensitivity, Hostility, Paranoid Ideation, and Psychoticism dimensions. These response sets appear to be independent of the quality of the translation, and cast doubt on the suitability of the bilingual technique as a means of assessing translation equivalence.

6.3.6 The Bilingual Technique

At the scale level, the bilingual test-retest results are shown in Table 8 (p. 78). The correlations range from .63 (Paranoid Ideation) to .82 (Interpersonal Sensitivity). These coefficients compare quite well with those obtained in the Zulu test-retest (see Table 3, p. 70). The higher correlations for the Paranoid Ideation and Psychoticism dimensions in the bilingual test-retest may be an artefact of the small sample size of the Zulu test-retest study, which may have produced unreliable results. Werner and Campbell (1970), viewing translation as a kind of paraphrase across two languages, argue that even the best translation can be no more similar to the original than a paraphrase in the original language. It is consequently unreasonable to expect cross-language correlations to be as high as ordinary test-retest correlations. At best, one might expect the cross-language correlations to approach those of within-language alternate forms across a comparable time delay. In view of this argument, the bilingual correlations obtained in the current study are particularly satisfactory.

Despite the satisfactory correlations, four of the symptom dimensions (Obsessive-Compulsive, Depression, Paranoid Ideation, and Psychoticism) produced significant discrepancies across the

two languages. For all four of these scales the English version scored higher. This is consistent with the language effect discussed above. The discrepancy was greatest in the case of the Paranoid Ideation and Psychoticism scales, which is consistent with an ethnic affirmation response set augmenting the language effect. The Hostility dimension was the only scale to score lower on the English version (although only marginally so), suggesting a social desirability effect which was strong enough to offset the language effect.

It is noteworthy that, in relation to the published norms (Derogatis, 1983), the scale scores for both language versions are extremely high for nonpatients and are more compatible with the norms for psychiatric patients. For example, the GSI means for the Zulu version (.97) and English version (1.07) correspond approximately to the 98th percentile for nonpatient males and the 50th percentile for outpatient males. This finding emphasises the importance of validating and renorming psychological tests that are to be used in different cultural groups, a point which will be addressed in Chapter 7.

In summary, at the scale level, which is likely to be of most interest to the clinician, the English and Zulu versions correlate well, suggesting that no major change has occurred as a result of the translation process. Mean score discrepancies across the two versions are consistent with bilingual response sets and are not necessarily indicative of translation errors.

The item-level bilingual test-retest results are shown in Table 9 (pp. 79-82). If equivalence has been achieved in the translation of an item, bilinguals' responses to the two versions, even allowing for the effects of the response sets discussed above, should display a positive correlational relationship. Negative correlations would suggest that an item had been mistranslated to the extent that it was measuring entirely different constructs in the two language versions.

None of the SCL-90-R items correlated negatively, providing support for the equivalence of the two versions. However, three items (numbers 17, 19, and 71) correlated close to zero and were examined for possible translation errors. The committee members were unable to identify any errors and concluded that these items had been accurately translated. An alternative explanation for the poor correlations is that they may reflect the comprehensibility of the two versions as applied to the sample, rather than the quality of the translation. Whereas the Zulu items may have been accurately translated and comprehensible, the three English items may have been particularly difficult for the respondents to understand. For example, item number 17: "Trembling" consists of a single word which may have been unfamiliar to many in the sample tested, and which offers no context or redundancy to assist the respondent in deducing its meaning. In contrast, the Zulu translation, *ukuqhaqhazela*, is a common Zulu expression which was well understood by the pretest sample.

Comparison of item means across the two versions is another way of assessing translation equivalence. If these means are comparable then "operational equivalence" (Prince & Mombour, 1967) has been achieved. In the context of the current study where it is apparent that response sets, which are independent of the quality of the translation, are influencing responses across the two languages, a conservative approach to interpreting the significance of item discrepancies seems warranted. Consequently, only those items with discrepancies significant at $p < .01$ were considered for possible revision by the committee.

The ranges of item means were quite similar for both versions (Zulu: .15 - 1.85; English: .30 - 1.90). The item mean discrepancies ranged from zero to .81. Out of the total of 90 items, 65 (72%) showed no significant discrepancy, providing further support for the adequacy of the translation. A further 11 items were significant at $p < .05$. The remaining 14 items

(16%) were significant at $p < .01$ or $p < .001$ and were considered by the committee.

The items in question are numbers 6, 10, 15, 18, 22, 24, 42, 51, 56, 57, 58, 73, 74, and 85. These items are distributed quite evenly across the nine symptom dimensions, three being on the Somatization dimension, two on each of the Obsessive-Compulsive, Depression, and Hostility dimensions, and one on each of the remaining scales. This is reassuring since items which may be problematic are not concentrated on any particular scale. In the case of ten of the items the discrepancy was in the direction of English higher, suggesting the influence of the dominant response sets (acquiescence and ethnic affirmation) discussed earlier.

No obvious translation errors could be identified by the committee. Unlike pretesting, where reasons for responses are recorded, the bilingual technique offers no insight into the reasons for discrepancies or how the discrepant items should be revised. If any changes were to have been made to the Zulu items, these would have had to be on a trial-and-error basis, with the bilingual technique being repeated until a satisfactory level of equivalence was achieved. Given the apparent distortions in the bilingual responses, and since considerable time and effort had previously been invested in ensuring that the items were accurately translated and comprehensible, such an approach seems unsound. Consequently, in the absence of clear evidence of translation errors, it was decided not to make further revisions to the Zulu SCL-90-R.

In summary, the psychometric evaluation of the Zulu SCL-90-R provides considerable support for the adequacy of the translation. Satisfactory measures of internal consistency and test-retest reliability suggest that the translation into Zulu had relatively little effect on the instrument's reliability. Scale-level factor analyses yielded a single large factor which accounted for most of the variance of both the Zulu and English

versions. This similarity of factor structure serves to demonstrate the fundamental comparability of the two versions. Due to retest effects and apparent bilingual response sets, the suitability of the bilingual technique as a means of assessing translation equivalence in the current study was called into question. It was consequently decided not to revise items based purely on the results of the bilingual technique. Finally, high mean scores of the Zulu subjects, relative to the published American norms for nonpatients (Derogatis, 1983), emphasise the need for the Zulu SCL-90-R to be subjected to criterion validation and renorming. This last point is addressed in the following chapter.

CHAPTER 7

PRELIMINARY VALIDATION STUDY

As emphasised by Derogatis (1983), validation is an ongoing process and is not achieved by means of a single definitive procedure. The results of predictive, concurrent, content, convergent, discriminant, and other types of validity studies contribute to the validation of the hypothetical constructs that the test is intended to operationalize. The importance of this process of validation cannot be overemphasised, particularly when one considers that psychological tests are used to guide professionals in making decisions which may have far-reaching implications for the lives of others.

With regard to translated tests, the author agrees with Butcher and Pancheri (1976) that the translator should assume some responsibility for demonstrating the criterion validity of the translated instrument. Accordingly, a preliminary validation study, relating to the concurrent validity of the Zulu SCL-90-R, was performed. The scores of a small sample of adult male psychiatric inpatients were compared with those of a sample of healthy Zulu men. The purpose of this study was to determine whether the Zulu translation of the SCL-90-R could discriminate between nonpatients and psychiatric inpatients. In addition, a cross-cultural comparison of SCL-90-R scores was conducted, the scores of the Zulu psychiatric inpatients and nonpatients being compared with the corresponding American norms in the SCL-90-R manual (Derogatis, 1983). If the Zulu scores were to differ significantly from the published norms, renorming of the Zulu SCL-90-R would be indicated.

7.1 METHOD

7.1.1 *Subjects*

The psychiatric sample (N = 23) consisted of male inpatients at two medium-term wards at Midlands Hospital Complex (Fort Napier). The inclusion criteria were that the patients be Zulu-speaking, cooperative, and capable of answering questions about themselves. The ward staff assisted in selecting patients according to these criteria. Informed consent was obtained from each of the patients, who were advised of the confidential nature of the research and that their decisions regarding whether or not to participate in the study would have no bearing on their subsequent treatment. The mean age of the sample was 30.8 years (SD = 9.7; range: 18 - 56). The mean number of years of formal schooling was 6.8 (SD = 2.8; range: 0 - 12). The mean number of days since admission to the hospital was 19 (SD = 12). All of the patients were receiving medication as part of their treatment. Eight of the patients (35%) had previous psychiatric admissions. Diagnostic information, based on hospital records, is shown in Table 10.

Table 10.

Diagnostic Categories of the Psychiatric Sample (N = 23)

Diagnostic Categories	Number of Patients
Major Depression	2
Manic Episode	3
Schizophrenia	3
Schizophreniform Disorder	1
Substance-Induced Psychotic Disorder	3
Psychotic Disorder NOS	4
Substance Abuse	3
Diagnosis Deferred	4

The nonpatient sample (N = 26) consisted of male workers employed at the Estates Division of the University of Natal. The mean age of the workers was 38.1 years (SD = 11.4; range: 22 - 59), somewhat higher than the psychiatric sample. With regard to education, this sample had a mean of 7 years of schooling (SD = 3.2, range: 0 - 12), which is comparable with the psychiatric sample. Based on the fact that the subjects were functioning at a high enough level to be full-time employees at the University, it was assumed that they were free of psychiatric illness. This assumption was not verified on an individual basis. The men had a variety of job descriptions, including driving, gardening, and painting. All had volunteered to participate in the study and understood its confidential nature.

7.1.2 *Procedure*

The Zulu SCL-90-R was administered verbally by trained Zulu research assistants, each subject being interviewed individually and privately. The assistants were instructed to spend some time discussing the SCL-90-R example, and to ensure that each subject understood what was required of him before commencing with the testing. Since the capacity of respondents to remember the five response categories during verbal administration should not be assumed, the categories were repeated for each item.

The data were analyzed by means of SPSS.

7.2 RESULTS

7.2.1 Concurrent Validity

7.2.1.1 *Comparison of Mean Scores*

The Zulu psychiatric inpatient and nonpatient scale scores are compared in Table 11. For all nine symptom dimensions and the three global scores, the levels of distress reported by the inpatients are significantly higher than those of the nonpatients.

Table 11.

Comparison of Mean Scores on the Zulu SCL-90-R for Psychiatric Inpatients and Nonpatients

	Patients (N = 23)		Nonpatients (N = 26)		Difference Between Means (Pt - Nonpt)	
Scale	Mean	SD	Mean	SD		Sig[a]
SOM	1.50	.94	.95	1.02	+.55	*
O-C	1.87	.99	1.25	.98	+.67	*
INT	1.61	1.08	.80	.90	+.81	**
DEP	1.88	.82	.76	.77	+1.12	***
ANX	2.02	1.07	.77	.93	+1.25	***
HOS	1.46	1.00	.85	.94	+.61	*
PHOB	1.60	1.04	.66	.83	+.94	**
PAR	1.78	1.12	.81	.84	+.97	**
PSY	1.96	.97	.53	.66	+1.43	***
GSI	1.77	.80	.84	.79	+.93	***
PSDI	3.12	.62	1.95	.71	+1.17	***
PST	49.65	18.79	33.12	25.37	+16.53	*

[a]Mann-Whitney U test *p<.05 **p<.01 ***p<.001

7.2.1.2 *Validity Coefficients*

The mean GSI score of the combined 49 cases is 1.27. Using this as a cutting score, it is possible to evaluate the efficiency of the Zulu SCL-90-R in distinguishing between the two groups. The classification of patients and nonpatients is shown in Table 12. The chi-square for the entire table (8.86) is significant at $p = .003$.

Table 12.

Classification of Zulu Patients and Nonpatients

Actual Samples	SCL-90-R Classification		Total
	Patients	Nonpatients	
Patients	16	7	23
Nonpatients	6	20	26
Total	22	27	49

Based on the figures shown in Table 12, the validity coefficients for the Zulu SCL-90-R are given in Table 13.

Table 13.

Validity Coefficients for the Zulu SCL-90-R

Sens	Spec	FNR	FPR	Class	BR
.70	.77	.30	.23	.73	.47

Notes. Sens = Sensitivity: proportion of patients correctly classified as patients. Spec = Specificity: proportion of nonpatients correctly classified as nonpatients. FNR = False Negative Rate: proportion of patients incorrectly classified as nonpatients (1 - Sens). FPR = False Positive Rate: proportion of nonpatients incorrectly classified as patients (1 - Spec). Class = Classification Rate: proportion of subjects correctly classified. BR = Base Rate: proportion of patients in total sample.

7.2.2 Cross-Cultural Comparison

7.2.2.1 *Comparison of Psychiatric Inpatients*

The scores of the Zulu and American psychiatric inpatients are compared in Table 14. Six of the symptom dimensions (Somatization, Anxiety, Hostility, Phobic Anxiety, Paranoid Ideation, and Psychoticism), as well as the GSI and Positive Symptom Distress Index (PSDI) were significantly higher for the Zulu patients.

Table 14.

Comparison of Mean SCL-90-R Scores for Zulu and American Psychiatric Inpatients

Scale	Zulu Inpatients (N = 23)	US Inpatients (N = 313)[a]	Difference (Zulu - US)	Significance[b]
SOM	1.50	.99	+.51	**
O-C	1.87	1.45	+.42	NS
INT	1.61	1.32	+.29	NS
DEP	1.88	1.74	+.14	NS
ANX	2.02	1.48	+.54	*
HOS	1.46	.94	+.52	*
PHOB	1.60	.96	+.64	**
PAR	1.78	1.26	+.52	*
PSY	1.96	1.11	+.85	***
GSI	1.77	1.30	+.47	**
PSDI	3.13	2.15	+.98	***
PST	49.65	50.03	-.38	NS

[a]Derogatis (1983) [b]Two-tailed t-test

*p<.05 **p<.01 ***p<.001

7.2.2.2 *Comparison of Nonpatients*

The scores of the Zulu and American nonpatients are shown in Table 15. All scores were substantially and significantly higher for the Zulu subjects.

Table 15.

Comparison of Mean SCL-90-R Scores for Zulu and American Nonpatients

Scale	Zulu Nonpatients (N = 26)	US Nonpatients (N = 974)[a]	Difference (Zulu - US)[b]
SOM	.95	.36	+.59
O-C	1.25	.39	+.86
INT	.80	.29	+.51
DEP	.76	.36	+.40
ANX	.77	.30	+.47
HOS	.84	.30	+.54
PHOB	.66	.13	+.53
PAR	.81	.34	+.47
PSY	.53	.14	+.39
GSI	.83	.31	+.52
PSDI	1.95	1.32	+.63
PST	33.11	19.29	+13.82

[a]Derogatis (1983)

[b]All differences significant at $p<.001$ (two-tailed t-test)

To facilitate comparison, the symptom profiles of all four samples, based on mean raw scores, are shown together in Figure 4.

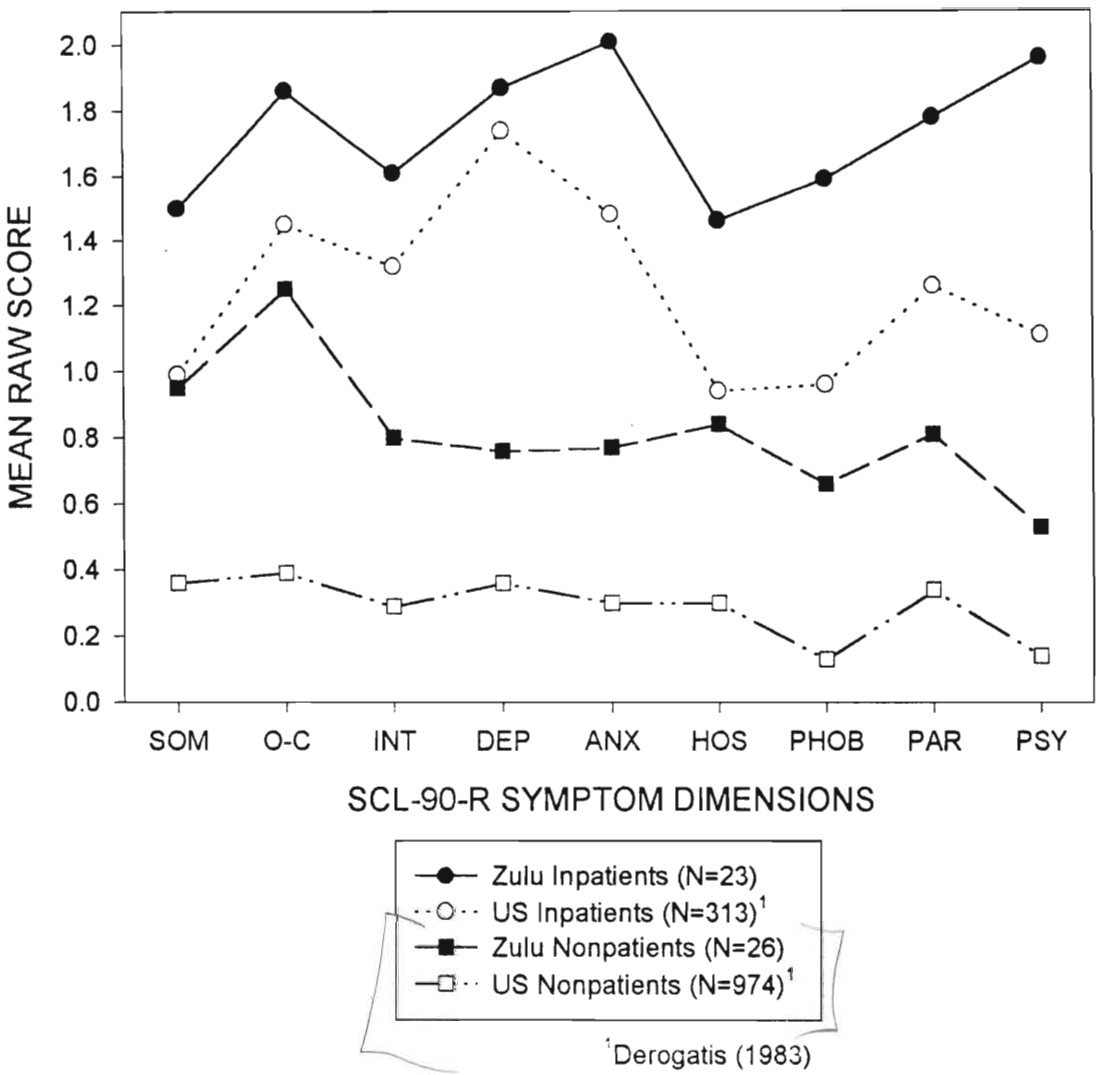

Figure 4 SCL-90-R Symptom dimension means for American and Zulu psychiatric inpatients and nonpatients

7.3 DISCUSSION

7.3.1 Concurrent Validity

7.3.1.1 *Comparison of Mean Scores*

The comparison of mean scores of the psychiatric patient and nonpatient samples on the Zulu SCL-90-R (see Table 11, p. 105) shows that the patients scored significantly higher on all of the symptom dimensions as well as the three global distress indices.

Of the symptom dimensions, the difference in scores was particularly substantial for the Depression, Anxiety, and Psychoticism scales. The high Psychoticism score for the patients may be a reflection of the composition of this sample. Consideration of the diagnostic information shown in Table 10 (p. 103) reveals that the psychiatric sample tends to be more representative of severe psychopathology, specifically the psychotic disorders, than milder forms of psychiatric illness. This may reflect the lack of psychiatric services available to the general population in South Africa, hospitalization being reserved for the most debilitated patients. Also, it is plausible that culturally-influenced help-seeking behaviours among Zulu people may result in only those with more severe symptomatology, causing extreme disruption to both the patient and those around them, being admitted to psychiatric facilities. Less disrupted individuals may be treated by traditional means and cared for within the community.

Regarding the global SCL-90-R measures, the higher scores of patients on the Positive Symptom Total (PST) indicates that they tend to report experiencing significantly more symptoms than do the nonpatients, while the PSDI scores indicate that the reported intensity of the symptoms is substantially higher for the patients. The significant difference in scores on the GSI supports the use of this measure for screening purposes.

7.3.1.2 *Validity Coefficients*

The validity coefficients for the Zulu SCL-90-R, using the GSI mean for the combined 49 cases (1.27) as a cutting score, are shown in Table 13 (p. 106). In the context of psychological testing, the sensitivity of 70% and specificity of 77% may be described as moderate, and suggest that the test may be validly utilised in screening for psychiatric illness. The SCL-90-R Manual (Derogatis, 1983) does not provide validity coefficients, and none of the published articles reviewed by the researcher give coefficients relating to the efficiency of the SCL-90-R in distinguishing between patients and nonpatients. It is consequently difficult to evaluate the validity of the Zulu SCL-90-R relative to other versions.

In the absence of more directly comparable studies, these results may be compared with those of a series of studies (Wetzler et al., 1989; Wetzler & Marlowe, 1993) investigating the diagnostic efficiency of individual scales of the SCL-90-R. Wetzler et al. (1989) examined the diagnostic efficiency of the SCL-90 Depression scale for the diagnosis of major depression. The test was administered to a group of patients diagnosed with major depression (n = 48) and a control group of nondepressed patients with diverse psychiatric disorders (n = 68). Describing the diagnostic efficiency of the Depression scale as "fairly good", Wetzler et al. (1989) report a sensitivity of .67, a specificity of .72, and a classification rate of .70.

In a subsequent study, Wetzler and Marlowe (1993) again assessed the diagnostic efficiency of the Depression scale in diagnosing depression, as well as that of the Paranoid Ideation and Psychoticism scales in diagnosing Psychosis. These scales were tested on a sample of 256 heterogeneous psychiatric patients (96 depressed, 38 manic, 77 psychotic). The Depression scale performed poorer than in the previous study (sensitivity = .56, Specificity = .67, Classification Rate = .61), while the Paranoid Ideation scale (sensitivity = .23, specificity = .74, classification

rate = .30) and Psychoticism scale (sensitivity = .36, specificity = .56, classification rate = .30) were insensitive to the diagnosis of psychosis.

Any comparisons between the results of the current study and those of Wetzler et al. (1989) and Wetzler and Marlowe (1993) should be drawn with caution. While the coefficients obtained in the above two studies are lower than those obtained in the current study, it may be argued that the challenges associated with making a differential diagnosis based on individual subscales are greater than those associated with distinguishing heterogeneous psychiatric patients from nonpatients based on a global score. Nevertheless, the higher coefficients of the current study are encouraging in terms of attempting to establish the validity of the Zulu SCL-90-R relative to other versions. Future studies should attempt to verify these results employing larger samples. The diagnostic efficiency of the Zulu SCL-90-R with regard to specific disorders should also be investigated, as well as its efficiency in screening for psychiatric illness in patients presenting for medical attention. Also, despite the unidimensionsal structure of the Zulu SCL-90-R indicated by the factor analysis of the responses of normal bilingual students (see Chapter 6), the possibility that diagnostic efficiency may be improved by utilizing symptom dimension scores (for example, via discriminant function analysis) should be investigated on larger samples.

An important consideration in screening for psychiatric illness is the effect of prevalence, or base rates, on the criterion validity of psychological tests. For a test of given sensitivity and specificity, the lower the prevalence of psychiatric illness in the population tested, the greater the problem of misclassification (Meehl & Rosen, 1955). In particular, lower base rates tend to result in increased numbers of false positives, healthy individuals who are misclassified as pathological. This problem is illustrated, based on a

hypothetical cohort of *N* = 1000 and using the moderate validity coefficients obtained for the Zulu SCL-90-R in the current study, in Table 16. With a prevalence rate of 30% the predictive value of a positive (PVP) of the test is 57%. At the lower prevalence rate of 10% the PVP drops to 25%. This means that three out of four positive screenings would be false positives.

Table 16.

Relationship between Prevalence and the Predictive Value of a Positive (PVP)

Test Sensitivity = .70
Test Specificity = .77

A. Prevalence = .30

	Test Classification		
REAL	Pos	Neg	
Pos	210	90	300
Neg	161	539	700
	371	629	1000

PVP = 210/371 = .57

B. Prevalence = .10

	Test Classification		
REAL	Pos	Neg	
Pos	70	30	100
Neg	207	693	900
	277	723	1000

PVP = 70/277 = .25

The predictive value of a negative (PVN) is not affected in the same way, remaining fairly high for both prevalence rates. The number of false negatives is thus relatively low (14% at .30 prevalence, and 4% at .10 prevalence).

Sequential screening is a method designed to cope with the problems of low base rates and moderately valid psychological tests (Derogatis & DellaPietra, 1994). This process involves two phases of screening. Phase I requires a test of at least moderate validity, the primary purpose being to eliminate a substantial proportion of those individuals who are free of disorder. This has the effect of raising the base rate of the disorder in the remaining cohort. In Phase II, a second test of at least equal sensitivity to the first is utilized. Since this

test is being applied to a cohort with a higher base rate of the disorder, there will be lower rates of false positive misclassification. Those individuals who screened positive could then be referred for more intensive individual examination. This process would allow for the relatively efficient utilization of available resources.

The results of this study, which due to the small samples employed should be seen as preliminary, suggest that the Zulu SCL-90-R may be validly utilised to distinguish between healthy individuals and those in need of psychiatric treatment.

7.3.2 Cross-Cultural Comparison

7.3.2.1 *Comparison of Psychiatric Inpatients*

The mean SCL-90-R scores for the Zulu and American psychiatric inpatients are compared in Table 14 (p. 107). The scores are higher for the Zulu patients on all nine of the symptom dimensions, and significantly so for six of the dimensions. The difference is particularly high in the case of the psychoticism dimension, which, as previously discussed, may reflect the composition of the Zulu sample.

Consideration of the global scores shows that the mean PST scores for the two groups are virtually identical at approximately 50, whereas the PSDI scores are substantially higher for the Zulu patients. This indicates that the generally higher scores of the Zulu patients do not reflect a larger *number* of symptoms being reported, but that this group tends to report experiencing higher levels of distress on those items which they do endorse. This may suggest higher levels of pathology in the Zulu sample, or that the Zulu patients tend to exhibit a more augmenting style of responding, or both.

As is apparent from Figure 4 (p. 109), the symptom profiles of the Zulu and American patients, with the noted exception of

the Psychoticism dimension, are more alike than dissimilar. Both groups exhibit high scores on the Obsessive-Compulsive, Depression, and Anxiety dimensions, and low scores on the Somatization, Hostility, and Phobic Anxiety dimensions. This cross-cultural similarity is promising with regard to the validity of the Zulu SCL-90-R, and suggests that the symptom constructs of the original may be generalizable to the Zulu version.

7.3.2.2 *Comparison of Nonpatients*

The scores of the Zulu and American nonpatients are compared in Table 15 (p. 108). On all measures, the Zulu group scored significantly and substantially higher than the Americans. These results indicate a cross-cultural difference in the baseline symptomatology of nonpatients, and strongly support the need for full-scale standardisation and renorming of the Zulu adaptation of the SCL-90-R. The score differences may reflect cross-cultural stylistic differences in reporting psychological distress. However, given South Africa's poor socio-economic conditions, inadequate health care, political violence, high rates of illiteracy, and other pathogenic conditions which primarily impact upon the country's black populations, it is possible that these differences may indicate higher levels of actual psychological distress in the Zulu group.

Notwithstanding the overall quantitative difference in scores, the symptom profiles of the two nonpatient groups (see Figure 4, p. 109) are broadly similar. For both groups, the dimensions of Obsessive-Compulsive and Somatization show relatively high scores, while Phobic Anxiety and Psychoticism scored the lowest. The Zulu profile is also similar to that for a community sample of Koreans (Kim, Kim, & Won, as cited in Noh & Avison, 1992). The Korean profile also showed a highest score for the Obsessive-Compulsive dimension (a mean raw score of approximately 1.20) and lowest scores for the Phobic Anxiety and Psychoticism dimensions (scores of approximately .52 and .65

respectively). This cross-cultural similarity of the profile patterns of nonpatient groups provides additional support for the construct validity of the Zulu SCL-90-R.

In summary, the Zulu SCL-90-R shows potential for the purpose of screening for psychiatric illness, at least among adult male Zulus. Having established this potential, it is important that more extensive validity studies be conducted to expand on these results. Similar symptom profiles for the American and Zulu samples provide some support for the construct validity of the Zulu SCL-90-R. However, significantly higher symptom levels of both the patient and nonpatient samples on the Zulu SCL-90-R as compared with the American norms, indicate the need for renorming of the Zulu SCL-90-R.

CONCLUSION

This study has attempted to produce a rigorously translated Zulu-language adaptation of the SCL-90-R, to assess the psychometric equivalence of this adaptation and the original SCL-90-R, and to investigate the validity of the instrument for the purpose of screening for psychiatric illness. The major findings of the study may be summarised as follows:

(1) Consultations with Zulu mental health professionals and pretesting with Zulu workers suggest that the Zulu adaptation of the SCL-90-R is acceptable and comprehensible to the target population.

(2) Comparison of bilinguals' scores on the Zulu and English versions of the SCL-90-R indicates generally satisfactory correspondence, although retest effects and apparent bilingual response sets undermine the validity of the bilingual technique as a means of assessing translation equivalence.

(3) Satisfactory internal consistency and test-retest reliability coefficients for the Zulu version suggest that the translation had little effect on the reliability of the SCL-90-R.

(4) Scale-level factor analyses of bilinguals' scores on the Zulu and English versions yielded similar structures for both versions, providing no evidence to suggest that a qualitative change had occurred as a result of the adaptation. The obtained factor structure suggests that the Zulu SCL-90-R is best utilised as a unidimensional measure of psychological distress, which is consistent with the majority of previous factor analytic studies of the SCL-90-R.

(5) A preliminary validation study yielded moderate validity coefficients for the Zulu SCL-90-R, suggesting that the instrument may be suitable for the purpose of screening for mental illness in the target population.

Suggestions for future research are discussed at appropriate points in Chapters 6 and 7. Having produced an apparently acceptable Zulu adaptation of the SCL-90-R, it is important that the work continue and that the psychometric properties and applicability of the instrument in local settings be systematically explored. This work should contribute towards the establishment of valid and reliable norms for various patient and nonpatient groups.

REFERENCES

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.), [DSM-IV]. Washington, DC: APA.

Anastasi, A. (1990). *Psychological Testing* (6th ed.). New York: Macmillan.

Ballenger, J.C., Burrows, G.D., Dupont, R.L., Lesser, I.M., Noyes, R., Pecknold, J.C., Rifkin, A., & Swinson, R.P. (1988). Alprazolam in panic disorder and agoraphobia: Results from a multicenter trail. I. Efficacy in short-term treatment. *Archives of General Psychiatry, 45*, 413-422.

Ben-Porath, Y.S. (1990). Cross-cultural assessment of personality: The case for replicatory factor analysis. In J.N. Butcher & C.D. Spielberger (eds.), *Advances in Personality Assessment* (Vol. 8). Hillsdale, NJ.: Lawrence Erlbaum Associates.

Berry, J.W. (1980). Introduction to methodology. In H.C. Triandis & J.W. Berry (eds.), *Handbook of Cross-Cultural Psychology: Methodology* (Vol. 2). Boston: Allyn & Bacon.

Berry, J.W., Poortinga, Y.H., Segall, M.H., & Dasen, P.R. (1992). *Cross-Cultural Psychology: Research and applications*. Cambridge: Cambridge University Press.

Bond, M.H. & Yang, K.S. (1982). Ethnic affirmation versus cross-cultural accommodation: The variable impact of questionnaire language on Chinese bilinguals from Hong Kong. *Journal of Cross-Cultural Psychology, 13*, 169-185.

Bontempo, R. (1993). Translation fidelity of psychological scales: An item response theory analysis of an

individualism-collectivism Scale. *Journal of Cross-Cultural Psychology, 24*, 149-166.

Bonynge, E.R. (1993). Unidimensionality of SCL-90-R scales in adult and adolescent crisis samples. *Journal of Clinical Psychology, 49*, 212-215.

Bravo, M., Canino, G.J., Rubio-Stipec, M., & Woodbury-Farina, M. (1991). A cross-cultural adaptation of a psychiatric epidemiologic instrument: The Diagnostic Interview Schedule's adaptation in Puerto Rico. *Culture, Medicine and Psychiatry, 15*, 1-18.

Brislin, R.W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1*, 185-216.

Brislin, R.W. (1976). Introduction. In R.W. Brislin (ed.), *Translation: Applications and research*. New York: Gardner.

Brislin, R.W. (1980). Translation and content analysis of oral and written materials. In H.C. Triandis & J.W. Berry (eds.), *Handbook of Cross-Cultural Psychology: Methodology* (Vol. 2). Boston: Allyn & Bacon.

Brislin, R.W. (1986). The wording and translation of research instruments. In W.J. Lonner & J.W. Berry (eds.), *Field Methods in Cross-Cultural Research*. Beverly Hills: Sage.

Brislin, R.W., Lonner, W.J., & Thorndike, R.M. (1973). *Cross-Cultural Research Methods*. New York: Wiley.

Brophy, C.J., Norvell, N.K., & Kiluk, D.J. (1988). An examination of the factor structure and convergent and discriminant validity of the SCL-90-R in an outpatient clinic population. *Journal of Personality Assessment, 52*, 334-340.

Buntting, B.G. (1988). *Comparison of the Present State Examination and the Structured Clinical Interview after Translation into Zulu*. Unpublished Masters Thesis, University of Natal, Durban.

Buntting, B.G. & Wessels, W.H. (1991). The 'present state' examination and the structured clinical interview in Zulu. *South African Medical Journal*, *79*, 96-98.

Butcher, J.N. (1982). Cross-cultural research methods in clinical psychology. In P.C. Kendall & J.N. Butcher (eds.), *Handbook of Research Methods in Clinical Psychology*. New York: Wiley.

Butcher, J.N. & Clark, L.A. (1979). Recent trends in cross-cultural MMPI research and application. In J.N. Butcher (ed.), *New Developments in the Use of the MMPI*. Minneapolis: University of Minnesota Press.

Butcher, J.N. & Pancheri, P. (1976). *Handbook of Cross-National MMPI Research*. Minneapolis: University of Minnesota Press.

Cameron, O.G., Thyrer, B.A., Nesse, R.M., & Curtis, G.C. (1986). Symptom profiles of patients with DSM-III anxiety disorders. *American Journal of Psychiatry*, *143*, 1132-1137.

Candell, G.L., & Hulin, C.L. (1986). Cross-language and cross-cultural comparisons in scale translations: Independent sources of information about item nonequivalence. *Journal of Cross-Cultural Psychology*, *17*, 417-440.

Carpenter, K.M. & Hittner, J.B. (1995). Dimensional characteristics of the SCL-90-R: Evaluation of gender differences in dually diagnosed inpatients. *Journal of Clinical Psychology*, 51, 383-390.

Carrington, P., Collings, G.H., Benson, H., Robinson, H., Wood, L.W., Lehrer, P.M., Woolfolk, R.L., & Cole, J. (1980). The use of meditation-relaxation techniques for the management of stress in a working population. *Journal of Occupational Medicine*, *22*, 221-231.

Cheung, F.M. (1985). Cross-cultural considerations for the translation and adaptation of the Chinese MMPI in Hong Kong. In J.N. Butcher & C.D. Spielberger (eds.), *Advances in Personality Assessment* (Vol. 4). Hillsdale, NJ.: Laurence Erlbaum Associates.

Church, A.T., Katigbak, M.S., & Castaneda, I. (1988). The effects of language of data collection on derived conceptions of healthy personality with Filipino bilinguals. *Journal of Cross-Cultural Psychology*, *19*, 178-192.

Clark, A. & Friedman, M.J. (1983). Factor structure and discriminant validity of the SCL-90 in a veteran psychiatric population. *Journal of Personality Assessment*, *47*, 396-404.

Clark, L.A. (1985). A consolidated version of the MMPI in Japan. In J.N. Butcher & C.D. Spielberger (eds.), *Advances in Personality Assessment* (Vol. 4). Hillsdale, NJ.: Lawrence Erlbaum Associates.

Comrey, A.L. (1988). Factor-analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology*, *56*, 754-761.

Cyr, J.J., & Atkinson, L. (1986). Selection of the appropriate number of factors: A case of questionable dimensions. *Psychological Reports*, *58*, 251-256.

Cyr, J.J., Doxey, N.C.S., & Vigna, C.M. (1988). Factorial composition of the SCL-90-R. *Journal of Social Behaviour and Personality, 3*, 245-252.

Derogatis, L.R. (1983). *SCL-90-R Administration, Scoring and Procedures Manual-II*. Towson, MD: Clinical Psychometric Research.

Derogatis, L.R. & Cleary, P.A. (1977a). Factorial invariance across gender for the primary symptom dimensions of the SCL-90. *British Journal of Social and Clinical Psychology, 16*, 347-356.

Derogatis, L.R. & Cleary, P.A. (1977b). Confirmation of the dimensional structure of the SCL-90: A study in construct validation. *Journal of Clinical Psychology, 33*, 981-989.

Derogatis, L.R. & DellaPietra, L. (1994). Psychological tests in screening for psychiatric disorder. In M.E. Maruish (ed.), *The Use of Psychological Testing for Treatment Planning and Outcome Assessment*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Derogatis, L.R. & Lazarus, L. (1994). SCL-90-R, Brief Symptom Inventory, and Matching Clinical Rating Scales. In M.E. Maruish (ed.), *The Use of Psychological Testing for Treatment Planning and Outcome Assessment*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Derogatis, L.R., Lipman, R.S., & Covi, L. (1973). SCL-90: An outpatient psychiatric rating scale - preliminary report. *Psychopharmacology Bulletin, 9*, 13-27.

Derogatis, L.R., Lipman, R.S., Rickels, K., Uhlenhuth, E.H., & Covi, L. (1974). The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. *Behavioral Science, 19*, 1-15.

Derogatis, L.R. & Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine, 13*, 595-605.

Derogatis, L.R., Meyer, J.K., & King, K.M. (1981). Psychopathology in individuals with sexual dysfunction. *American Journal of Psychiatry, 138*, 757-763.

Derogatis, L.R., Rickels, K., & Rock, A. (1976). The SCL-90 and the MMPI: A step in the validation of a new self-report scale. *British Journal of Psychiatry, 128*, 280-289.

Derogatis, L.R., & Wise, T.N. (1989). *Anxiety and Depressive Disorders in the Medical Patient*. Washington, D.C.: American Psychiatric Press.

Derogatis, L.R., Yevzeroff, H., & Wittelsberger, B. (1975). Social class, psychological disorder, and the nature of the psychopathologic indicator. *Journal of Consulting and Clinical Psychology, 43*, 183-191.

Desjarlais, R., Eisenberg, L., Good, B., & Kleinman, A. (1995). *World Mental Health: Problems and priorities in low-income countries*. New York: Oxford University Press.

Dinning, W.D. & Evans, R.G. (1977). Discriminant and convergent validity of the SCL-90 in psychiatric inpatients. *Journal of Personality Assessment, 41*, 304-310.

Draguns, J.G. (1982). Methodology in cross-cultural psychology. In I. Al-Issa (ed.), *Culture and Psychopathology*. Baltimore, MD.: University Park Press.

Drennan, G. (1992). *Aspects of Translation in Psychological and Psychiatric Cross-Cultural Research*. Unpublished M.A. thesis, University of Cape Town.

Dunnigan, T., McNall, M., & Mortimer, J.T. (1993). The problem of metaphorical nonequivalence in cross-cultural survey research: Comparing the mental health statuses of Hmong refugee and general population adolescents. *Journal of Cross-Cultural Psychology, 24*, 344-365.

Ebigbo, P.O. (1982). Development of a culture specific (Nigeria) screening scale of somatic complaints indicating psychiatric disturbance. *Culture, Medicine and Psychiatry, 6*, 29-43.

Edwards, S.D., Cheetham, R.W.S., Majozi, E., & Lasich, A.J. (1985). Zulu culture-bound psychiatric syndromes. In S.D. Edwards (ed.), *Some Indigenous South African Views on Illness and Healing*. KwaDlangezwa: University of Zululand.

Edwards, S.D., Grobbelaar, L.M., Nene, L.M., Makunga, N.V., Kunene, S.T., & Sibaya, P.T. (1985). Reactions to illness and concepts of mental illness among representative samples of rural, urban and university educated black people. In S.D. Edwards (ed.), *Some Indigenous South African Views on Illness and Healing*. KwaDlangezwa: University of Zululand.

Ervin, S.M. (1964). Language and TAT content in bilinguals. *Journal of Abnormal and Social Psychology, 68*, 500-507.

Evenson, R.C., & Holland, R.A., Metha, S., & Yasin, F. (1980). Factor analysis of the Symptom Checklist-90. *Psychological Reports, 46*, 695-699.

Flaherty, J.A., Gaviria, F.M., Pathak, D., Mitchell, T., Wintrob, R., Richman, J.A., & Birz, S. (1988). Developing instruments for cross-cultural psychiatric research. *Journal of Nervous and Mental Disease, 176*, 257-263.

Freeman, M. (1992). *Providing Mental Health Care For All in South Africa: Structure and strategy*. Johannesburg: Centre for Health Policy, University of the Witwatersrand.

Garner, W.R. (1960). Rating scales, discriminability, and information transmission. *Psychological Review, 67*, 343-352.

Geisinger, K.F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment, 6*, 304-312.

Gillis, L.S., Elk, R., Ben-Arie, O., & Teggin, A. (1982). The Present State Examination: Experiences with Xhosa-speaking psychiatric patients. *British Journal of Psychiatry, 141*, 143-147.

Gordon, L.V., & Kikuchi, A. (1966). American personality tests in cross-cultural research: a caution. *Journal of Social Psychology, 69*, 179-183.

Guilford, J.P. (1954). *Psychometric Methods* (2nd ed.). New York: McGraw-Hill.

Hafkenscheid, A. (1993). Psychometric evaluation of the Symptom Checklist (SCL-90) in psychiatric inpatients. *Personality and Individual Differences, 14*, 751-756.

Hambleton, R.K., Swaminathan, H., & Rogers, H.J. (1991). *Fundamentals of Item Response Theory*. Newbury Park: Sage.

Hoffmann, N.G. & Overall, P.B. (1978). Factor structure of the SCL-90 in a psychiatric population. *Journal of Consulting and Clinical Psychology, 46*, 1187-1191.

Holcomb, W.R., Adams, N.A., & Ponder, H.M. (1983). Factor structure of the Symptom Checklist-90 with acute psychiatric inpatients. *Journal of Consulting and Clinical Psychology*, *4*, 535-538.

Horowitz, L.M., Rosenberg, S.E., Baer, B.A., Ureno, G., & Villasenor, V.S. (1988). Inventory of interpersonal problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology*, *56*, 885-892.

Horowitz, M.J., Wilner, N., Kaltreider, N., & Alvarez, W. (1980). Signs and symptoms of posttraumatic stress disorder. *Archives of General Psychiatry*, *37*, 85-92.

Hui, C.H., & Triandis, H.C. (1985). Measurement in cross-cultural psychology: a review and comparison of strategies. *Journal of Cross-Cultural Psychology*, *16*, 131-152.

Hulin, C.L. (1987). A psychometric theory of evaluation of items and scale translations: fidelity across languages. *Journal of Cross-Cultural Psychology*, *18*, 115-142.

Irvine, S.H. (1986). Cross-cultural assessment: From practice to theory. In W.J. Lonner & J.W. Berry (eds), *Field Methods in Cross-Cultural Research*. Beverly Hills: Sage.

Irvine, S.H. & Carroll, W.K. (1980). Testing and assessment across cultures: Issues in methodology and theory. In H.C. Triandis & J.W. Berry (eds.), *Handbook of Cross-Cultural Psychology: Methodology* (Vol. 2). Boston: Allyn & Bacon.

Jamison, R.N., Rock, D.L., & Parris, W.C.V. (1988). Empirically derived Symptom Checklist 90 subgroups of chronic pain patients: A cluster analysis. *Journal of Behavioral Medicine*, *11*, 147-158.

John, V.N. (1996). *Use of the General Health Questionnaire (GHQ) in a Zulu-speaking setting: An assessment of translation, reliability, and some validity issues*. Unpublished M.A. thesis, University of Natal, Pietermaritzburg.

Kleinman, A. (1987). Anthropology and psychiatry: The role of culture in cross-cultural research on illness. *British Journal of Psychiatry, 151*, 447-454.

Koeter, M.W.J. (1992). Validity of the GHQ and SCL anxiety and depression scales: A comparative study. *Journal of Affective Disorders, 24*, 271-280.

Kortmann, F. (1987). Problems in communication in transcultural psychiatry: The self reporting questionnaire in Ethiopia. *Acta Psychiatrica Scandinavica, 75*, 563-570.

Kortmann, F., & Ten Horn, S. (1988). Comprehension and motivation in responses to a psychiatric screening instrument: Validity of the SRQ in Ethiopia. *British Journal of Psychiatry, 153*, 95-101.

Lees-Haley, P.R. (1989). Malingering emotional distress on the SCL-90-R: Toxic exposure and cancerphobia. *Psychological Reports, 65*, 1203-1208.

Le Roux, A.G. (1973). Psychopathology in Bantu culture. *South African Medical Journal, 47*, 2077-2083.

Lezak, M.D. (1995). *Neuropsychological Assessment* (3rd ed.). New York: Oxford University Press.

Lissitz, R.W. & Green, S.B. (1975). Effect of the number of scale points on reliability: A Monte Carlo approach. *Journal of Applied Psychology, 60*, 10-13.

Malgady, R.G., Rogler, L.H., & Constantino, G. (1987). Ethnocultural and linguistic bias in mental health evaluation of Hispanics. *American Psychologist, 42*, 228-234.

Marin, G., Triandis, H.C., Betancourt, H., & Kashima, Y. (1983). Ethnic affirmation versus social desirability: Explaining discrepancies in bilinguals' responses to a questionnaire. *Journal of Cross-Cultural Psychology, 14*, 173-186.

Mauer, K.F., & Retief, A.I. (1985). Metodologiese probleme in die sielkunde. *South African Journal of Philosophy, 4*, 19-24.

Mazmanian, D., Mendonca, J.D., Holden, R.R., & Dufton, B. (1987). Psychopathology and response styles in the SCL-90 responses of acutely distressed persons. *Journal of Psychopathology and Behavioral Assessment, 9*, 135-148.

Meehl, P.E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin, 52*, 194-216.

Michelson, C.L. (1991). *Trauma, Levels of Distress, and Coping, Amongst Displacees in the Pietermaritzburg Area.* Unpublished M.Soc.Sc. thesis, University of Natal, Pietermaritzburg.

Miller, T., & Swartz, L. (1992). Psychology and epidemiology: an uncomfortable alliance? *South African Journal of Psychology, 22*, 52-58.

Ngubane, H. (1977). *Body and Mind in Zulu Medicine: An ethnography of health and disease in Nyuswa-Zulu thought and practice.* London: Academic Press.

Noh, S. & Avison, W.R. (1992). Assessing psychopathology in Korean immigrants: Some preliminary results on the SCL-90. *Canadian Journal of Psychiatry, 37*, 640-645.

Nunnally, J.C. (1978). *Psychometric Theory*, (2nd ed.). New York: McGraw-Hill.

O'Donnell, W.E., DeSoto, C.B., & Reynolds, D.M. (1984). A cognitive deficit subscale of the SCL-90-R. *Journal of Clinical Psychology, 40*, 241-246.

Orley, J. & Wing, J.K. (1979). Psychiatric disorders in two African villages. *Archives of General Psychiatry, 36*, 513-520.

Pareek, U. & Rao, T.V. (1980). Cross-cultural surveys and interviewing. In H.C. Triandis & J.W. Berry (eds), *Handbook of Cross-Cultural Psychology: Methodology* (Vol. 2). Boston: Allyn & Bacon.

Peltz, M. & Merskey, H. (1982). A description of the psychological effects of chronic painful lesions. *Pain, 14*, 293-301.

Peveler, R.C. & Fairburn, C.G. (1990). Measurement of neurotic symptoms by self-report questionnaire: Validity of the SCL-90-R. *Psychological Medicine, 20*, 873-879.

Platt, J.J., Steer, R.A., Ranieri, W.F., & Metzger, D.S. (1989). Differences in the Symptom Check List-90 profiles of black and white methadone patients. *Journal of Clinical Psychology, 45*, 342-345.

Poortinga, Y.H. (1989). Equivalence of cross-cultural data: an overview of basic issues. *International Journal of Psychology, 24*, 737-756.

Prince, R. & Mombour, W. (1967). A technique for improving linguistic equivalence in cross-cultural surveys. *International Journal of Social Psychiatry, 13*, 229-237.

Prusoff, B.A., Weissman, M.M., Klerman, G.L., & Rounsaville, B.J. (1980). Research diagnostic criteria subtypes of depression: Their role as predictors of differential response to psychotherapy and drug treatment. *Archives of General Psychiatry, 37*, 791-801.

Ravaris, C.L., Robinson, D.S., Ives, J.O., Nies, A., & Bartlett, D. (1980). Phenelzine and amitriptyline in the treatment of depression. *Archives of General Psychiatry, 37*, 1075-1080.

Retief, A. (1988). *Method and Theory in Cross-Cultural Psychological Assessment*. Pretoria: Human Sciences Research Council.

Rief, W. & Fichter, M. (1992). The Symptom Check List SCL-90-R and its ability to discriminate between dysthymia, anxiety disorders, and anorexia nervosa. *Psychopathology, 25*, 128-138.

Rogler, L.H. (1989). The meaning of culturally sensitive research in mental health. *American Journal of Psychiatry, 146*, 296-303.

Roskin, M. & Dasberg, H. (1983). On the validity of the Symptom Check List-90 (SCL-90): A comparison of diagnostic self-ratings in general practice patients and 'normals', based on the Hebrew version. *International Journal of Social Psychiatry, 29*, 225-230.

Schwartzwald, J., Weisenberg, M., & Solomon, Z. (1991). Factor invariance of SCL-90-R: The case of combat stress reaction. *Psychological Assessment, 3*, 385-390.

Sechrest, L., Fay, T.L., & Zaidi, S.M.H. (1972). Problems of translation in cross-cultural research. *Journal of Cross-Cultural Psychology, 3*, 41-56.

Shapiro, D.A., & Firth, J. (1987). Prescriptive versus exploratory psychotherapy: Outcomes of the Sheffield Psychotherapy Project. *British Journal of Psychiatry, 151*, 790-799.

Shaw, M. & Wright, J. (1967). *Scales for the Measurement of Attitudes*. New York: McGraw-Hill.

Simoes, M. & Binder, J. (1980). A socio-psychiatric field study among Portuguese emigrants in Switzerland. *Social Psychiatry, 15*, 1-7.

Sperber, A.D., Devellis, R.F., & Boehlecke, B. (1994). Cross-cultural translation: Methodology and validation. *Journal of Cross-Cultural Psychology, 25*, 501-524.

Steer, R.A., Clark, D.A., & Ranieri, W.F. (1994). Symptom dimensions of the SCL-90-R: A test of the tripartite model of anxiety and depression. *Journal of Personality Assessment, 62*, 525-536.

Steer, R.A. (1982). Modal profile analyses of the SCL-90-R for seven mental disorders. *Multivariate Experimental Clinical Research, 6*, 55-67.

Strauman, T.J. & Wetzler, S. (1992). The factor structure of SCL-90 and MCMI scale scores: Within-measure and interbattery analyses. *Multivariate Behavioral Research, 27*, 1-20.

Subedar, H. (1993). *Screening for Mental Disorders in the Mahlabatini Ward in the Ngcolosi District of the Valley of*

a Thousand Hills, Natal. Unpublished MCur Thesis, University of Natal, Durban.

Takeuchi, D.T., Kuo, H.-S., Kim, K., & Leaf, P.J. (1989). Psychiatric symptom dimensions among Asian Americans and Native Hawaiians: An analysis of the Symptom Checklist. *Journal of Community Psychology, 17*, 319-329.

Taylor, T.R. & Boeyens, J.C.A. (1991). The comparability of the scores of blacks and whites on the South African Personality Questionnaire: An exploratory study. *South African Journal of Psychology, 21*, 1-11.

Triandis, H.C. (1972). *The Analysis of Subjective Culture*. New York: Wiley.

Triandis, H.C. (1976). Approaches toward minimising translation. In R.W. Brislin (ed.), *Translation: Applications and research*. New York: Gardner.

Triandis, H.C. (1980). Introduction to Handbook of Cross-Cultural Psychology. In H.C. Triandis & W.W. Lambert (eds.), *Handbook of Cross-Cultural Psychology: Perspectives* (Vol. 1). Boston: Allyn & Bacon.

Tyson, G.A., Doctor, E.A., & Mentis, M. (1988). A psycholinguistic perspective on bilinguals' discrepant questionnaire responses. *Journal of Cross-Cultural Psychology, 19*, 413-426.

Uzzell, B., & Oler, J. (1986). Chronic low-level mercury exposure and neuropsychological functioning. *Journal of Clinical and Experimental Neuropsychology, 8*, 581-593.

Velicer, W.F., & Stevensen, J.F. (1978). The relation between item format and the structure of the Eysenck Personality Inventory. *Applied Psychological Measurement, 2*, 293-304.

Wakefield, J.C. (1992). The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist, 47*, 373-388.

Werner, O., & Campbell, D.T. (1970). Translating, working through interpreters, and the problem of decentering. In R. Naroll & R. Cohen (eds.), *A Handbook of Method in Cross-Cultural Anthropology*. New York: Natural History Press.

Westermeyer, J., Vang, T.F., & Neider, J. (1983). Migration and mental health among Hmong refugees: Association of pre- and postmigration factors with self-rating scales. *Journal of Nervous and Mental Disease, 171*, 92-96.

Wetzler, S., Kahn, R., Strauman, T.J., & Dubro, A. (1989). Diagnosis of major depression by self-report. *Journal of Personality Assessment, 53*, 22-30.

Wetzler, S. & Marlowe, D.B. (1993). The diagnosis and assessment of depression, mania, and psychosis by self-report. *Journal of Personality Assessment, 60*, 1-31.

Whorf, B.L. (1956). *Language, Thought and Reality*. New York: Wiley.

Wilson, J.H., Taylor, P.J., & Robertson, G. (1985). The validity of the SCL-90 in a sample of British men remanded to prison for psychiatric reports. *British Journal of Psychiatry, 147*, 400-403.

Woody, G.E., Luborsky, L., McLellan, T., O'Brien, C.P., Beck, A.T., Blain, J., Herman, I., & Hole, A. (1983). Psychotherapy for opiate addicts. *Archives of General Psychiatry, 40*, 639-645.

Yang, K.S. & Bond, M.H. (1980). Ethnic affirmation by Chinese bilinguals. *Journal of Cross-Cultural Psychology*, *11*, 411-425.

APPENDIX A: GUIDELINES FOR TRANSLATORS

Your task is to translate each item of the SCL-90-R into Zulu so that the original meaning is retained while the language used is likely to be understandable to a wide range of Zulu-speaking people of various backgrounds (educational, socio-economic, etc.). The SCL-90-R is intended for use with adults and adolescents (from age 13 up).

The following guidelines are intended to assist you in considering the various kinds of equivalence that you should attempt to achieve.

1. Vocabulary equivalence

This involves equivalence of the actual words used. While a good dictionary is a valuable resource, translation based purely on dictionary consultation is not sufficient to achieve vocabulary equivalence. The language contained in dictionaries is often not the language used by the people for whom the translated material is intended, and may consequently be misunderstood by test respondents.

There may be English terms for which Zulu equivalents do not exist. A possible solution is the use of a short description in Zulu in order to convey a concept which may be expressed by a single word in English. Differences in the length of materials should, however, be kept within reasonable limits. Where possible, try to keep the length of each sentence to less than 16 words.

2. Idiomatic equivalence

You will notice that the SCL-90-R contains some idiomatic expressions (e.g., "feeling blue") which may not be literally translated into Zulu. These items may be replaced by a phrase which conveys the meaning of the expression, or by an appropriate equivalent Zulu idiom if such exists. The addition

of an explanation, in parentheses, to clarify items containing figurative expressions may also be considered.

3. Grammatical-syntactical equivalence
Zulu and English differ in their grammars and syntaxes, and these differences may present problems for you as the translator. As a general rule, the effective transmission of *meaning* should take precedence over excessive concern with grammatical and syntactical considerations which could lead to long-winded and confusing language.

4. Conceptual and experiential equivalence
Travelling on subways (underground trains), is an example of an experience which would be foreign to Zulu people. While it may be necessary to eliminate such an item as culturally inappropriate, your task is to ascertain what the item was intended to measure in the original culture (e.g., phobic anxiety) and to find an equivalent which would be meaningful to Zulu individuals.

You are requested to rate each item according to the difficulty encountered:
1 = relatively easy to translate, equivalent meaning achieved
2 = some difficulty in translation, equivalence may be questionable
3 = considerable difficulty in translation, equivalence is questionable

You are also requested to make brief notes of particular difficulties encountered and decisions you made in dealing with these difficulties.

Thank you for your assistance.

APPENDIX B: ZULU TRANSLATION OF THE SCL-90-R

The following pages contain a copy of the Zulu translation of the SCL-90-R. The SCL-90-R is fully protected by copyright and may not be reproduced in any manner without the prior written permission of the author, Dr L.R. Derogatis, c/o NCS Assessments, 5605 Green Circle Drive, Minnetonka, MN 55343, USA.

IMININGWANE EQONDENE NAWE

1. IGAMA: _____

2. IMINYAKA YAKHO: ____

3. UBULILI:
   (i) ISILISA
   (ii) ISIFAZANE

4. IBANGA LOKUGCINA OWALIPHASA ESIKOLENI: _____

5. UMSEBENZI OWENZAYO: _____

6. ISIMO SOMSHADO:
   (i) USHADILE/UGANILE
   (ii) AWUSHADILE/AWUGANILE

7. NAMHLANJE NGUMHLAKA: _____

IMIYALELO

Ngezansi kunohla lwezinkinga ezike zikhungathe abantu kwesinye isikhathi. Uyacelwa ukuba ufunde inkinga ngayinye ngokucophelela, bese ukekelezela inombolo eyodwa echaza kangcono ngenkinga ekukhathazile noma ekuhluphile EZINSUKWINI EZIYISIKHOMBISA EZEDLULILE UHLANGANISE NOLWANAMUHLANJE. Kekelezela inombolo eyodwa kuleyo naleyo nkinga ungeqi neyodwa inkinga ohleni olunikeziwe. Uma kwenzeka uguqula umqondo, cima impendulo oyikhethe kuqala ngokucophelela. Funda isibonelo ngezansi ngaphambi kokuthi uqale, uma unemibuzo uvumelekile/uyacelwa ukuba ubuze ngayo.

ISIBONELO KUKUKHATHAZE KANGAKANANI LOKHU OKULANDELAYO:	AKUKAZE	KANCANE	MAPHAKATHI NJE	KAKHUDLWANA	KAKHULU IMPELA
1. UBUHLUNGU BOMZIMBA	0	1	2	(3)	4

KUKUKHATHAZE KANGAKANANI LOKHU OKULANDELAYO:	AKUKAZE	KANCANE	MAPHAKATHI NJE	KAKHUDLWANA	KAKHULU IMPELA
1. UKUPHATHWA YIKHANDA	0	1	2	3	4
2. UKWESABA NOMA UVALO	0	1	2	3	4
3. UKUBA NEMICABANGO EMIBI ENGAPHELI ENGQONDWENI YAKHO	0	1	2	3	4
4. ISIYEZI NOMA INZULULWANE	0	1	2	3	4
5. UKUPHELELWA UTHANDO NOMA UBUMNANDI NGOKOCANSI	0	1	2	3	4
6. UKUZIZWA UNOKUGXEKA ABANYE	0	1	2	3	4

KUKUKHATHAZE KANGAKANANI LOKHU OKULANDELAYO:	AKUKAZE	KANCANE	MAPHAKATHI NJE	KAKHUDLWANA	KAKHULU IMPELA
7. UMQONDO WOKUTHI OMUNYE UMUNTU ANGALAWULA IMICABANGO YAKHO	0	1	2	3	4
8. UKUZIZWA UNOKUGXEKA ABANYE NGENINGI LEZINKINGA ZAKHO	0	1	2	3	4
9. INKINGA YOKUKHOHLWA	0	1	2	3	4
10. UKUKHATHAZEKA UBUDLABHA NOMA UKUNGAQIKELELI	0	1	2	3	4
11. UKUZIZWA UNOKUTHUKUTHELA NOMA UKUCASUKA KALULA	0	1	2	3	4
12. IZINHLUNGU ESIFUBENI	0	1	2	3	4
13. UKUZIZWA UNOKWESABA EZINDAWENI EZIVULEKILE NOMA EZITALADINI	0	1	2	3	4
14. UKUZIZWA UPHELELWA UMDLANDLA NOMA UZWE UMZIMBA UPHANSI	0	1	2	3	4
15. UKUFIKELWA IMICABANGO YOKUZIBULALA	0	1	2	3	4
16. UKUZWA AMAZWI ABANYE ABANTU ABANGAWEZWA	0	1	2	3	4
17. UKUQHAQHAZELA	0	1	2	3	4
18. UKUZWA SENGATHI ININGI LABANTU LINGETHENJWE	0	1	2	3	4
19. UKUNGAKUTHANDI UKUDLA	0	1	2	3	4
20. UKUKHALA KALULA	0	1	2	3	4
21. UKUZIZWA UNAMAHLONI NOMA UNGAKHULULEKILE UMA UNABANTU BOBULILI OBAHLUKILE	0	1	2	3	4
22. UKUZIZWA UVIMBELEKILE NOMA UBAMBEKILE ESIMENI ESITHIZE	0	1	2	3	4
23. UKWESABA OKUSHESHAYO NGAPHANDLE KWESIZATHU	0	1	2	3	4
24. UKUHLUTHUKA KWENHLIZIYO NOMA UKUBA NOLAKA ONGAKWAZI UKUKUGWEMA	0	1	2	3	4
25. UKUZIZWA WESABA UKUPHUMA ENDLINI UWEDWA	0	1	2	3	4
26. UKUZIGXEKA NGEZINTO EZENZEKA KUWE	0	1	2	3	4
27. IZINHLUNGU EQOLO	0	1	2	3	4
28. UKUZIZWA UVIMBELEKILE EKWENZENI IZINTO UZIQEDE	0	1	2	3	4
29. UKUZIZWA UNESIZUNGU	0	1	2	3	4
30. UKUZIZWA UDABUKILE NOMA UDANGELE	0	1	2	3	4
31. UKUKHATHAZEKA KAKHULU NGEZINTO	0	1	2	3	4
32. UKUZIZWA UNGENALO UTHANDO LWEZINTO	0	1	2	3	4
33. UKUZIZWA UNOKWESABA	0	1	2	3	4
34. UKUZIZWA UPHATHEKA KABI KALULA EMPHEFUMULWENI	0	1	2	3	4
35. ABANYE ABANTU BAZI IMICABANGO ESENHLIZIYWENI YAKHO	0	1	2	3	4

KUKUKHATHAZE KANGAKANANI LOKHU OKULANDELAYO:	AKUKAZE	KANCANE	MAPHAKATHI NJE	KAKHUDLWANA	KAKHULU IMPELA
36. UKUZWA SENGATHI ABANYE ABAKUQONDI NOMA ABAKUZWELI	0	1	2	3	4
37. UKUZWA SENGATHI ABANTU ABAKWENAMELI NOMA ABAKUTHANDI	0	1	2	3	4
38. UKWENZA IZINTO KANCANE KAKHULU UKUZE UBE NESIQINISEKO SOKWENZISISEKA	0	1	2	3	4
39. UKUSHAYA KWENHLIZIYO KAKHULU NOMA NGOKUSHESHA	0	1	2	3	4
40. UKUCANUZELA KWENHLIZIYO NOMA UKUPHENDUKELWA YISISU	0	1	2	3	4
41. UKUZINYEZA	0	1	2	3	4
42. UBUHLUNGU BEZICUBU ZOMZIMBA NOMA AMAMASELA	0	1	2	3	4
43. UKUZWA SENGATHI ABANYE BAKUBHEKILE NOMA BAKHULUMA NGAWE	0	1	2	3	4
44. UKUBA NENKINGA YOKUNGAFIKELWA UBUTHONGO	0	1	2	3	4
45. UKUBA NESIDINGO SOKUHLOLA UPHINDE UHLOLE OKWENZAYO	0	1	2	3	4
46. UKUBA NOBUNZIMA EKUTHATHENI IZINQUMO	0	1	2	3	4
47. UKUZIZWA UNOKWESABA UKUHAMBA NGAMABHASI, NOMA UKUHAMBA NGEZITIMELA	0	1	2	3	4
48. UKUPHEFUMULA KANZIMA	0	1	2	3	4
49. UKUBA NEZIKHAWU ZOKUSHISA NOMA ZOKUBANDA	0	1	2	3	4
50. UKUGWEMA IZINTO EZITHILE, IZINDAWO NOMA UKWENZA IZINTO EZITHILE NGOBA ZIKWESABISA	0	1	2	3	4
51. UKUMELWA YINGQONDO	0	1	2	3	4
52. UKUNGABI NAMIZWA NOMA UKUBA NENKWANTSHU KWEZINGXENYE ZOMZIMBA	0	1	2	3	4
53. UKUBA NESIGAXA EMPHINJENI WAKHO	0	1	2	3	4
54. UKUPHELELWA YITHEMBA NGEKUSASA	0	1	2	3	4
55. UKUBA NENKINGA EKUGXILENI ENTWENI OYENZAYO	0	1	2	3	4
56. UKUZIZWA UPHELELWA AMANDLA NOMA UKUBA NTEKENTEKE EZINGXENYENI ZOMZIMBA WAKHO	0	1	2	3	4
57. UKUZIZWA UNGAKHULULEKILE EMZIMBENI NOMA UXAKANISEKILE EMPHEFUMULWENI	0	1	2	3	4
58. UKUZIZWA USINDWA YIZINGALO NOMA IMILENZE	0	1	2	3	4
59. IMICABANGO YOKUFA NOMA YOKUTHI UYAFA	0	1	2	3	4
60. UKUDLA NGOKWEQISA	0	1	2	3	4
61. UKUZIZWA UNGAKHULULEKILE UMA ABANTU BEKUBUKA NOMA BEKHULUMA NGAWE	0	1	2	3	4
62. UKUBA NEMICABANGO ENGESIYONA EYAKHO, EKUNGATHI IFAKWA ABANYE ABANTU EMQONDWENI WAKHO	0	1	2	3	4
63. UKUBA NELUKULUKU LOKUSHAYA, LOKULIMAZA NOMA LOKWENZAKALISA OMUNYE	0	1	2	3	4

KUKUKHATHAZE KANGAKANANI LOKHU OKULANDELAYO:	AKUKAZE	KANCANE	MAPHAKATHI NJE	KAKHUDLWANA	KAKHULU IMPELA
64. UKUPHAPHAMA EKUSENI KAKHULU BESE WEHLULEKA UKUZUMEKA FUTHI	0	1	2	3	4
65. UKUBA NELUKULUKU LOKUPHINDAPHINDA IZENZO EZIFANAYO NAKUBA USUKE WAZENZA NJENGOKUGEZA IZANDLA, UKUBALA IZINTO NOMA UKUHLOLA UKUTHI UMNYANGO UKHIYIWE	0	1	2	3	4
66. UKUQWASHA	0	1	2	3	4
67. UKUBA NELUKULUKU LOKWEPHULA NOMA UKUPHAHLAZA IZINTO	0	1	2	3	4
68. UKUBA NEMIBONO NOMA IZINKOLELO ABANYE ABANGAHAMBISANI NAZO	0	1	2	3	4
69. UKUBA NOKUZINYEZA LAPHO UPHAKATHI KWABANYE	0	1	2	3	4
70. UKUZIZWA UNGAKHULULEKILE LAPHO UPHAKATHI KWABANTU ABANINGI, NJENGASEZITOLO NASEMABHAYISIKOBHO	0	1	2	3	4
71. UKUZWA SENGATHI YONKE INTO OYENZAYO INGUMZABALAZO	0	1	2	3	4
72. IZIWOMBE ZOVALO NOMA UKWETHUKA	0	1	2	3	4
73. UKUZIZWA UNGAKHULULEKILE NGOKUDLA NOMA NGOKUPHUZA PHAKATHI KWABANTU	0	1	2	3	4
74. UKUHLALA NJALO UHILIZISANA NABANYE NGAMAZWI	0	1	2	3	4
75. UKUZIZWA UNOVALO UMA USELE WEDWA	0	1	2	3	4
76. ABANYE ABAKUNCOMI NGOKUFANELE NGEMPUMELELO YAKHO	0	1	2	3	4
77. UKUZIZWA UNESIZUNGU NOMA UNABANYE ABANTU	0	1	2	3	4
78. UKUZIZWA UNGAPHUMULE, UNGENASINQE, KANGANGOBA UNGAKWAZI NOKUHLALA UZOTHE	0	1	2	3	4
79. UKUZIZWA UNGENTO YALUTHO	0	1	2	3	4
80. UKUZWA SENGATHI KUKHONA INTO EMBI EZOKWENZEKA KUWE	0	1	2	3	4
81. UKUTHETHA NOMA UKUPHOSA IZINTO	0	1	2	3	4
82. UKUZWA WESABA UKUTHI UZOQULEKA PHAMBI KWABANTU	0	1	2	3	4
83. UKUBA NOKUZWA UKUTHI ABANTU BAZOKUXHAPHAZA UMA UBAVUMELA	0	1	2	3	4
84. UKUBA NEMICABANGO EKUHLUPHA KAKHULU NGEZOCANSI	0	1	2	3	4
85. UKUBA NOMQONDO WOKUTHI UFANELE UJEZISWE NGEZONO ZAKHO	0	1	2	3	4
86. IMICABANGO NEMIFANEKISO EYETHUSAYO	0	1	2	3	4
87. UKUBA NOMQONDO WOKUTHI KUKHONA INTO ENKULU EYONAKELE EMZIMBENI WAKHO	0	1	2	3	4
88. AWUKAZE UZIZWE USONDELE KOMUNYE UMUNTU	0	1	2	3	4
89. UKUZIZWA UNECALA	0	1	2	3	4
90. UMCABANGO WOKUTHI KUKHONA OKUNGALUNGILE ENGQONDWENI YAKHO	0	1	2	3	4