# EVALUATING THE CONSTRUCT VALIDITY OF THE

# KIDSCREEN-52 QUALITY OF LIFE QUESTIONNAIRE

# WITHIN A SOUTH AFRICAN CONTEXT UTILISING

# EXPLORATORY FACTOR ANALYSIS: INITIAL VALIDATION

**Naiema Taliep**

A mini-thesis submitted in partial fulfilment of the requirements for

the degree of M.A. (Research) Psychology in the Department of Psychology,

University of the Western Cape
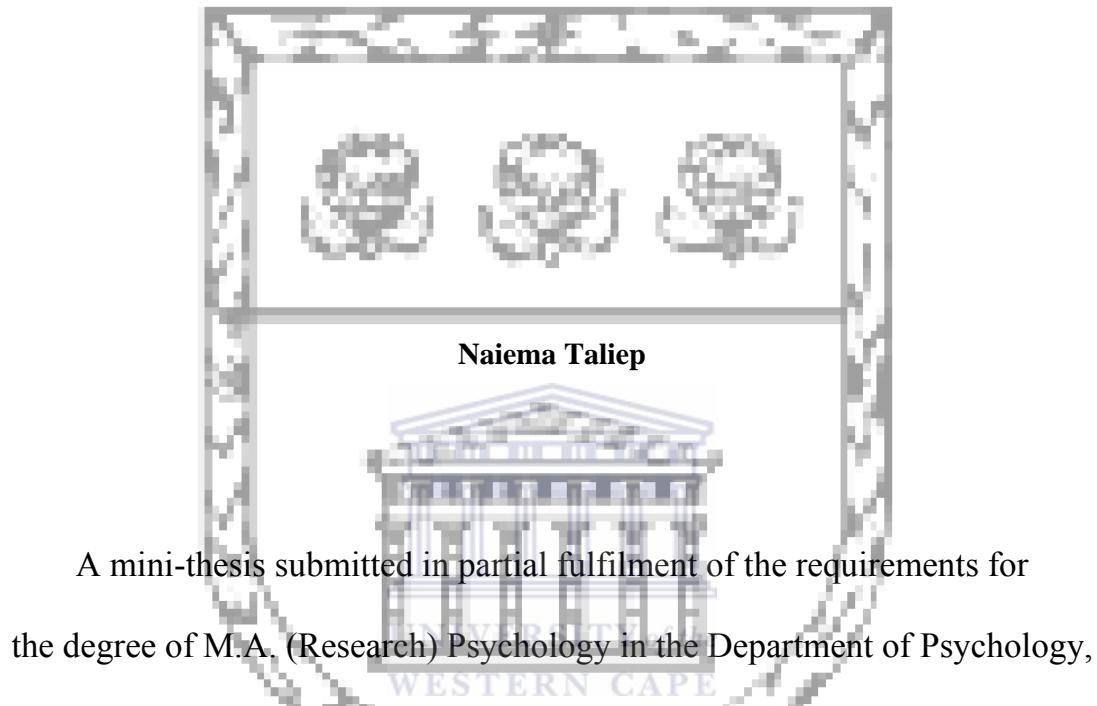
**Supervisor: Maria Florence**

**September 2010**

**Keywords:** Assessment, Quality of Life, KIDSCREEN-52, children and adolescents, South Africa, subjective perception, secondary data analysis, construct validity, reliability, exploratory factor analysis.

# ABSTRACT

The absence of a suitable measure to assess the health related quality of life (HRQoL) of children and adolescents in South Africa, led to the use of the KIDSCREEN-52 questionnaire in this study. This questionnaire was developed and standardised in Europe and it assesses quality of life from the subjective perception of children and adolescents in terms of their physical, mental and social well-being. The use of assessment measures with different populations or in different milieu from the original standardisation context necessitates establishing the validity of such measures for the new utilization context. Thus, the aim of this quantitative study is specifically directed at examining the construct validity of the KIDSCREEN-52 within a South African context. Accordingly, this study is located within the theoretical framework of construct validation theory. Data for this Secondary Data Analysis study was drawn from the "Impact of Hope and Exposure to Community Violence on children's perception of Well-being" study. The primary study employed stratified interval criterion sampling to select 565 grade 9 learners, aged 14-18 from six public schools. The dataset for the current study comprised all participants ($N$=565) of the broader study. As the initial step in validation of the KIDSCREEN-52 within South Africa, the current study examined the factor structure of the KIDSCREEN-52 within this context by means of exploratory factor analysis using principal component analysis with oblimin rotations. It also assessed the internal consistency reliability of each of the scales using Cronbach's alpha. Exploratory factor analysis revealed the same 10 factors as identified by previous European studies with some deviation in the last two factors, which warrants further examination. Internal consistency of the measure was shown to be acceptable, with Cronbach's alpha values ranging from 0.76 to 0.81 for the 10 scales.

**DECLARATION**

The author hereby declare that the following final research report, " Evaluating the Construct Validity of the KIDSCREEN-52 Quality of Life questionnaire within a South African context utilizing Exploratory Factor Analysis: Initial validation" is of her own work and all the sources she has used or quoted have been indicated and acknowledged by means of complete references. This study used secondary data analysis and fully acknowledges the investigators of the original study.

Naiema Taliep                                                                    September 2010

Signed: ...........……………………

UNIVERSITY *of the*
WESTERN CAPE

# ACKNOLEDGEMENTS

# TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

Page

UNIVERSITY of the
WESTERN CAPE

## LIST OF FORMULAS

Page

**CHAPTER ONE**

**INTRODUCTION**

**1.1 Background**

Over the last number of years, the development of measures that are geared towards positive mental health and well-being or Quality of Life (QoL) has become the vogue internationally in both social sciences and health. Notwithstanding the prominence QoL instruments have gained, Robitail et al. (2006) assert that there is still a dire need for measures that would permit Health Related Quality of Life (HRQoL) assessments at an international level. It is precisely the lack of sufficient data on the subjective health of children and adolescents in different European countries which led to the simultaneous development of the KIDSCREEN-52 questionnaire within 13 European countries (Ravens-Sieberer, Schmidt et al., 2007).

Similar to the international trend, a review of South African literature verifies that there is a lack of information regarding the HRQoL of children and adolescents within this country. The absence of a suitable assessment measure in a particular country compels researchers to either develop a new instrument, to adapt an existing instrument, or to "export" an existing, almost always Western measure, to the non-Western world (Van de Vijver & Rothman, 2004) and establish its psychometric properties in the new context.

To address the dire need of adequate data which provides information about the subjective HRQoL of South African children and adolescents, Savahl, Isaacs, September and Koch (2009) constructed a questionnaire to investigate the impact of hope and exposure to community violence on children's perception of well-being (HECVW) by using the KIDSCREEN-52 in conjunction with the Children's Hope Scale and the Recent Exposure to

Violence Scale (REVS). Participants for this study were drawn from six public schools within areas specified by the South African government as comprising key nodal areas in terms of crime (high, medium and low violence neighbourhoods) in the Western Cape. The present study represents an initial attempt to establish the psychometric properties of the KIDSCREEN-52 in a South African context. It can therefore be regarded as an extension of the HECVW study. The focus of this study specifically centres on establishing the construct validity and reliability of the KIDSCREEN-52 within a South African context.

The fact that the bulk of currently available generic HRQoL assessment measures have been constructed within a particular country or culture is a cause for concern, as what is regarded as important to HRQoL in one country may not be as important in other countries or cultures (Herdman et al., 2002). The utilisation of such measures leaves test users with little empirical conviction that the tests they use are valid, reliable and free from bias. It is thus important to establish the psychometric properties of such measures within the new utilisation context. Moreover, the disreputable roots of psychological testing in South Africa served as the impetus to ethical and social considerations in the development and administration of psychological measures. Cultural appropriateness of psychological tests and their usage came into the spotlight in South Africa with the promulgation of the new Employment Equity Act 55 of 1998 (Van de Vijver & Rothman, 2004) which stipulates that psychological testing or assessments are forbidden unless the test or assessment being used has been shown to be valid and reliable and conforms to ethics of fairness and bias (Government Gazette, 1998). Robitail et al. (2007) concur that it is vital to determine the validity of measures for use in different populations. Since the construct validity of the KIDSCREEN-52 has not been established within a South African context, it warrants the need for the current study.

The construct, QoL itself has been applied over a broad expanse of health related problems such as drug abuse, oncology and mental health (Claes, Van Hove, van Loon, Vandevelde & Schalock, 2009). Despite its prominence, however, there remains a lack of consensus as to what constitutes QoL. Whilst Fallowfield (2009) maintains that this ubiquitous concept has various philosophical, political and health-related dimensions, O'Connor (1993) states that definitions of QoL in the health context are, for the most part, vague or lacking. Thus, it is a widely diffused pervasive notion which makes it difficult to define. However, HRQoL is generally conceptualised as a multi-dimensional psychological concept that encompasses functioning and well-being in the physical, social and mental or emotional dimensions of life (Fallowfield, 2009; Hays, 2003).
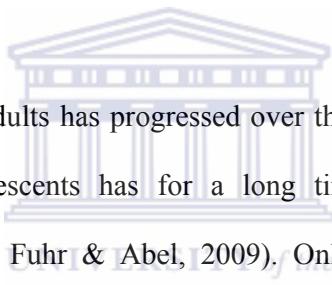
Even though QoL research in adults has progressed over the past number of years, HRQoL research in children and adolescents has for a long time been neglected and under-investigated (Michel, Bisegger, Fuhr & Abel, 2009). Only 13% of all HRQoL research publications are related to children, and of 320 publications identified, only 9% refers to testing of assessment instruments (Ravens-Sieberer & Bullinger, 1998). A number of instruments have been developed to assess the QoL in children and adolescents, but instrument content was mostly based on proxy data from parents. Moreover, Ravens-Sieberer and Bullinger (1998) state that the most commonly utilised mode of assessment are parent and proxy assessments concerning child well-being, with more than 90% of studies showing that self-report measures are lacking. Yet, in reviewing conceptual, methodological and regulatory issues in the assessment of HRQoL, Matza, Swensen, Flood, Secnik and Leidy (2004) confirm that HRQoL is subjective and should therefore be assessed from the patient's perspective whenever possible.

HRQoL among children and adolescents is thus a recent domain (Ravens-Sieberer et al., 2005) which has been plagued by various methodological challenges. Concerns that emerged include the reliability of data obtained from children and adolescents, the age-appropriateness of the measure in relation to their literacy level, the use of proxy data, and the domains to be covered in the measure. To begin with, even though children are generally regarded as unreliable respondents (Claes et al., 2009), recent research provides evidence that children can reliably and validly self-report their HRQoL when age-appropriate measures are used (Lundqvist, Rugland, Clench-Aas, Bartonova & Hofoss, 2010; Myant & Williams, 2005; Riley, 2004; Ravens-Sieberer et al., 2005; Varni, Limbers & Burwinkle, 2007). In addition, their emotional development, cognitive capacity as well as reading skills should also be taken into account (Ravens-Siberer et al., 2005). Secondly, the use of indirect or proxy data is a greatly debated issue and test users are warned that proxy data reflect a doubtful substitute which should only be utilised in exceptional cases (Ravens-Sieberer et al., 2006). As far as the domains are concerned, QoL dimensions differ with the developmental phases (Bullinger, Schmidt, Petersen & Ravens-Sieberer, 2006). Thus, developmental changes should be taken into consideration when assessing the HRQoL of children and adolescents. In view of this, adapting and utilising adult measures to assess children and adolescents is therefore also questionable (Ravens-Sieberer et al., 2006).

The KIDSCREEN-52 has been developed as a standardised instrument that can be applied in paediatric (ailing or disabled) and healthy populations with equal relevance to assess the subjective HRQoL of children and adolescents (Ravens-Sieberer et al., 2005). It was constructed based on very stringent procedures and on the opinions of the target population. A Delphi panel of experts was assembled and together they specified the necessary components of a QoL assessment instrument for children and adolescents (Rajmil et al.,

2004; Ravens-Sieberer et al., 2006). The KIDSCREEN-52 therefore consists of items that have been demonstrated to be generically age and culture appropriate and comparable (Ravens-Sieberer et al., 2006). It is a cross-culturally applicable measure developed following various cross-cultural approaches. It was developed simultaneously in a number of European countries and contains country specific as well as multicultural aspects (Ravens-Sieberer et al., 2005).  However, these cross-cultural validations primarily occurred in Europe and the suitability for its cross-cultural use beyond Europe needs to be established.

## 1.2 Rationale

The value of QoL questionnaires in research rests squarely upon their validity and physicians cannot interpret QoL measures until the instruments being evaluated are adequately established (Muldoon, Barger, Flory & Manuck, 1998). Claes and associates (2009) advise that any QOL assessment instrument should address construct related questions. Failure to assess construct validity is tantamount to turning a blind eye to corrupting elements entrenched in measures (e.g. measurement error) which can seriously jeopardise the conclusions drawn in a study (O'Leary-Kelly & Vokurka, 1998). The psychometric properties of the KIDSCREEN-52 were assessed using the Classical Test Theory approach, Rasch analysis and structural equation modeling (SEM) (Ravens-Sieberer, Gosch et al., 2008). Robitail and associates (2007) assessed the validity of the KIDSCREEN-27, a shorter version of the KIDSCREEN-52, by testing its multi-dimensional structure, the unidimensionality of its five dimensions and its internal consistency, and the existence of any cross-cultural differences in item functioning. They also performed Exploratory Factor Analysis (EFA), using principal component analysis (PCA), on the sample as a whole. Even though the KIDSCREEN-52 is a well-validated measure and its construct validity has already been established (Ravens-Sieberer, Schmidt et al., 2007), validity is not a conclusive feature

of a test, but relative to every specific purpose for which the test is used (Jooste, 2001). Thus, when a test is used for a purpose beyond the original standardisation validation conditions, then the validity of that test for the new utilisation conditions, should be determined again (Jooste, 2001). Accordingly, since the construct validity of the KIDSCREEN-52 has not been established in the South African context, it informs the need for this study.

By establishing the construct validity of the KIDSCREEN-52 within a South African context, the current study provides an initial description of the psychometric properties of this measure in this context. Psychometric authorities regard construct validity as the chief criterion for the validation of psychological measures (Cronbach & Meehl, 1955; Messick, 1989). The construct is usually embedded in some theory concerning behaviour, and the theory clarifies the meaning of the construct, how it is connected to other constructs, and how it is connected to specific observable behaviours (Ghiselli, Campbell & Zedeck, 1981). Moreover, construct validation is not once-off; it is a process which is ongoing and continuous, a process that examines the adequacy of a test in relation to theory (Strauss & Smith, 2009). This study is therefore located within the theoretical framework of Construct Validation Theory.

Furthermore, construct validation is regarded as a multifaceted process which, according to O'Leary-Kelly and Vokurka (1998) should embrace the establishment of the degree to which the empirical indicators measure the construct. They further state that this entails the empirical assessment of the adequacy of a measure and necessitates the establishment of validity (the extent to which a scale or set of measures actually measures the characteristic we intend to measure) and reliability (the extent to which a measure produces consistent results).

Reliability is an important feature of an assessment instrument as unreliability detracts from validity (Pesudous, Burr, Harley & Elliot, 2007).

## 1.3 Aims

The overall aim of this study is therefore to assess the construct validity of the KIDSCREEN-52 within a South African context by means of Secondary Data. The specific aims of this study are as follows:

Research aim 1: To examine the factor structure of KIDSCREEN-52 within a South African context to determine whether this study replicates the 10-dimensional structure identified by previous European studies.

Research aim 2: To assess the internal consistency reliability of each of the sub-scales of the KIDSCREEN-52 using Cronbach's alpha.

## 1.4 Significance of study

The psychometric properties are vital in evaluating the usefulness of a psychological measure, and since a measure is only valid for the specific purpose for which it was designed, its use in a different context necessitates the re-establishment of its validity in the new context (Foxcroft & Roodt, 2005). This study will therefore augment the main study, i.e. the HECVW study, by assessing the construct validity of the KIDSCREEN-52 in a South African context. In addition, considering the current paucity of adequate QoL measures for children and adolescents (Ravens-Sieberer et al., 2005) in South Africa as well as abroad, assessing the validity of existing measures for use with different populations can narrow this gap.

**1.5 Conclusion**

This introduction provided a framework of the current study and its contribution to the broader study within which it is located. It also provided a brief definition of the central construct of this study, and highlighted pertinent issues pertaining to measuring the construct HRQoL. It concluded with the aims and significance of the study.

Chapter two will explore the theoretical framework by looking at construct validity within psychometric theory as well as the link between the acceptable operationalisation and conceptualisation of the construct of import and the establishment of construct validity. Following this, chapter three will provide a comprehensive review of the available literature on HRQoL in children and adolescents; chapter four will look at the methodology employed in conducting the study and chapter five will comprise the results. Finally, chapter six will provide a discussion of the results, limitations and recommendations as well as the final conclusion.
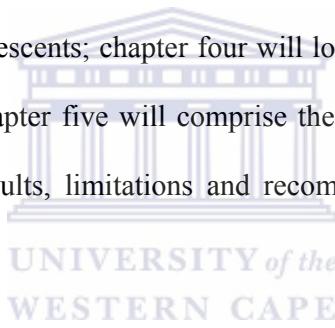
# CHAPTER TWO

# THEORETICAL DISCUSSION

## 2.1 Introduction

The current study is encased in the theoretical framework of construct validation theory. A fundamental prerequisite of any assessment measure is that it must be valid. Validity is defined as a unified, but faceted, evolving property and validation is regarded as scientific theory (Messick, 1989). In test validation, the attribute or quality about which we make statements or draw inferences when interpreting a test is a construct (Cronbach & Meehl, 1955). Construct validity is therefore defined as the degree to which an assessment measure can be interpreted as a meaningful measure of a particular characteristic or construct (Onwuegbuzie, Daniel & Collins, 2009). In establishing construct validity, we therefore demonstrate that the results are consonant to the definition, i.e. the psychological nature of the construct. It is thus a process where the theory and the measure are assessed at the same time. Furthermore, construct validity of a test is to be evaluated by examining the complete body of evidence presented, along with what is postulated about the measure in the context of this evidence (Cronbach & Meehl, 1955). Thus, given the significance of construct validity to the assessment of HRQoL, it is vital to have a clear overview of the various facets relating to construct validation.

This chapter therefore serves as a basic framework for contextualising the validation of the KIDSCREEN-52 within a South African context. It highlights the pivotal role of construct validation theory in the validation of the KIDSCREEN-52. In order to demonstrate this, this chapter will first provide a brief history on the evolving nature of the concept validity, the

process of establishing validity evidence, determining construct validity, and threats to construct validity.

## 2.2 Evolving approaches to validity

In order to get a more comprehensive understanding of the validation process, we need to first explore how the concept validity developed in psychometric theory. The term 'validity' generally refers to the extent to which a scale or set of measures actually measures the characteristics or dimensions we intend to measure (Walsh & Betz, 2001) or accurately represent the concept being measured. However, the actual definition of validity has evolved over the years. According to Goodwin (2000), four general periods in relation to definitions of validity can be identified. In the first period (1920 to 1950) test validity focused on the test itself since a test was regarded as valid for whatever it correlated with. Kaplan, Bush and Berry (1976) state that prior to 1950, most social scientists only considered criterion related validity (i.e. how well scores on the measure correlate with the criterion of interest) as well as content related validity (i.e. the extent to which the items on a measure represent the content being measured). However, authors on validity during this period showed a great deal of dissatisfaction with conventional notions of validity and introduced new terms and ideas, which only caused more confusion (Cronbach & Meehl, 1955). Developing a measure of QoL at that time, for example, would have posed a dilemma since a precise definition of such constructs was lacking. Investigators then concluded that no clear criteria existed for most of the social measures being developed which eventually led to the establishment of more explicit foundations for the assessment of construct validity (Kaplan et al., 1976).

In the second period (1950-1970) the validity of a test for a specific purpose, with a specific population, within a specific setting, became significant (Goodwin, 2000). The term construct

validity was then introduced in 1955 by Cronbach and Meehl in one of the most important detailed expositions on measurement ever published (Anastasi & Urbina, 1997; Sechrest, 2005). During this time, another ground-breaking development in validation theory and method occurred with the introduction of the multitrait- multimethod matrix introduced by Campbell and Fiske in 1950 (Goodwin, 2000). This meant that multiple approaches to construct validation became important in an inclusive study of the construct validity of a measure. They advocated the need for convergent validity (i.e. illustrating the extent to which a measure is related to other tests of the same construct) and discriminant validity (i.e. illustrating the degree to which a measure is not related to tests of different constructs) (Goodwin, 2000). The third major period emerged during the 1980s with Messick and Cronbach proposing modified definitions of validity that call attention to the inferences to be drawn from test scores (Goodwin, 2000).

The current view, which marks the fourth period in the evolution of definitions of validity, is Messick's (1989) notion that validity investigations must embrace the study of the social consequences of test use. He defined validity as the overall assessment of the extent to which empirical evidence and theoretical justifications bolsters the trustworthiness and appropriateness of interpretations and actions based on instrument scores. A recent definition advanced by the American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCME) (1999) concurs that validity refers to the extent to which evidence and theory bolster the interpretations of test scores implied by postulated uses of tests. This implies that given a set of specific questions we want the KIDSCREEN-52 to help answer, how useful or appropriate (i.e. valid) the information provided by the test scores is (Ghiselli et al., 1981).

Hence, the concept validity also relates to the theoretical and applied usefulness of a measure (Walsh & Betz, 2001). The usefulness of a particular test according to Walsh and Betz (2001) hinges on our ability to make deductions about people or environments from the test scores, and the kind of inferences we aspire to draw depends on the purposes of that test. However, they state that it is the extent and precision of possible inferences to other real-world phenomena that serves as evidence for the validity of a test.

## 2.3 Establishing validity evidence

The process of validation entails the systematic collection of evidence to provide a sound scientific justification for the inferences that are intended to be drawn from scores obtained from a measure (AERA, APA & NCME, 1999). Traditionally, in validation studies, researchers set out to establish one or more of three distinct types of validity, i.e. content-related validity, criterion related validity or construct related validity (Onwuegbuzie et al., 2009). Currently, there is a consensual view that validity is a unitary concept (AERA, APA, & NCME, 1999; Goodwin, 2000; Messick, 1989; Onwuegbuzie et al., 2009) and construct validity embraces all forms of validity. Since validity is regarded as a unitary concept instead of three distinct types of validity, the present focus is on types of validity evidence (AERA, APA & NCME, 1999). Content and criterion validation procedures are therefore among the different sources of information that contribute to the definition and the understanding of the construct/s assessed by a measure (Anastasi & Urbina, 1997), i.e. they provide validity evidence. Messick (1998) regards them as complementary forms of evidence which needs to be incorporated into the overall assessment of construct validity.

Loevinger's (as cited in Messick, 1989) rationale in support of construct validity is predicated on the notion that construct validity is the whole of validity. Furthermore, he states that

content, concurrent and predictive validities are all basically ad hoc from a scientific perspective whereas, construct validity is comprehensive in nature. For that reason, Messick (1989) broadened his theory of construct related validity to all sources of evidence that supports the meaning and interpretation of scores from a measure as well as actions based on such interpretations. He therefore draws a distinction between the "evidential" and the "consequential" basis of test validity.

This distinction denotes a difference between factual and value-laden aspects or social consequences of testing. The one facet of construct validity entails the trustworthiness of score meaning and its interpretation. The establishment of this aspect of test validity is what Messick (1989) calls the evidential basis of test validity. The evidential or factual aspects of construct validity relate either to logical stipulations or empirical factors, or a combination of both, within a normal scientific paradigm (Orton, 1998). Messick (1989) emphasises that in psychological measurement, inferences are derived from scores, which in a general sense refers to any coding or summarisation of observed consistencies on an observation procedure or any assessment device. Validity, therefore, at all times denotes the degree to which theoretical rationales and empirical evidence sustain the adequacy and suitability of inferences and actions derived from test scores (Messick, 1989).

The consequential aspects of construct validity pertain to the moral and social implications of test use. According to Orton (1998), even though test users have always (to various degrees) conceded that their work is value-laden, Messick incorporates this normative element into his concept of validity – a notion conventionally associated with 'truth'. Emphasis is placed on the inadvertent consequences of specific uses of tests that could negatively impact on individuals or members of certain ethnic groups or populations from diverse experiential

backgrounds (Anastasi & Urbina, 1997). Messick (1998) states that score interpretations have social consequences which essentially contribute to score meaning and therefore to construct validity. The social consequences of score interpretations comprise the value implications attached to the construct label, to the broader theory within which the construct is located, and even broader ideologies that restrict and colour theory construction (Messick, 1989).

A number of authors disagree with Messick's notion of consequential validity. Sechrest (2005) voices serious reservations regarding the applicability of consequential validity. Markus (1998) argues that Messick's theory of validity contains an unfinished synthesis, whereas Anastasi and Urbina (1997) state that linking the empirical, statistically supported procedures of establishing validity with an assessment of the social and ethical consequences of specific test uses only obscures and obfuscates the solution. In defence of his view, Messick (1989) cites Cronbach's line of reasoning that validity does not only refer to the accuracy of inferences drawn from test scores, but also to the assessment of the appropriateness, meaningfulness and utility of score inferences, which inexorably involves value judgments.

In this respect, it is important to note that tests are not just developed to draw inferences about individuals, but they are also utilised to make important decisions as in the case of high stakes tests. Orton (1998) states that Messick's distinction between consequential and evidential aspects of test validity is a safety measure against treating person's as things, especially in the face of high stakes testing decisions. Moreover, if consequences are not part of the validation process, numerous sources of invalidity will continue to remain unexposed to the disadvantage of the discipline as well as the practice of psychological measurement (Messick, 1998). For example, if the KIDSCREEN-52 is used to assess the impact of public

health interventions within a given population, any decision based on scores obtained can have grave consequences for individuals if it impacts decisions taken on a certain course of action.

**2.4 Threats to construct validity**

There are a number of threats to construct validity which fall into two major classes, i.e construct irrelevant variance and construct under-representation. The first major threat to construct validity, construct irrelevant variance, refers to unrelated components that sneak into measurement and contaminate it (Messick, 1989). These components are extraneous or unrelated to the construct of interest, but they produce reliable variance in test scores which is irrelevant to the construct (Messick, 1989).

Another major threat to construct validity, which according to Messick (1989) clouds the meaning and interpretation of test scores, is construct under-representation. This relates to the inability of tests to adequately tap all the features of a construct due to certain deficiencies, such as not including important dimensions of the construct (Messick, 1989). When constructing a test, we garner features of the construct of interest in accordance with our definition of the construct which we intend to measure. If we left out certain intrinsic features that we should have included, it constrains the score meaning and interpretation. According to the AERA, APA and NCME (1999), it denotes a limited meaning of test scores since the measure does not adequately sample certain types of contents, engage certain psychological processes, or obtain some ways of responding that are incorporated by the intended construct. Moreover, Singleton, Straits and Straits (1993) maintain that the issue of validity cannot generally be separated from larger theoretical interests such as what the nature of your concept is, what it means and whether your operational definition aptly represents this

meaning or something else. If a test of QoL, for example, purports to assess QoL, but in reality measures self-esteem then it is not a valid measure of quality of life. The trait being measured by a specific test can only be defined by means of an assessment of the objective sources of information and empirical operations used in establishing its validity (Anastasi & Urbina, 1997). Assessing the validity of a measure, therefore, begs a precise knowledge of the psychological domain under consideration along with a clear operational definition of each characteristic being measured (Hammond, 2006).

To guard against construct under-representation, the development of the KIDSCREEN-52 assessment measure was constructed based on very stringent procedures and on the opinions of the target population. Firstly, researchers identified that most existing measures were based on expert opinion regarding what constitutes HRQoL, whilst the opinions and perceptions of healthy children were seldom included in such measures (Detmar, Bruil, Ravens-Sieberer, Gosch & Bisegger, 2006). In order to identify the gaps in previous conceptualisation of the HRQoL construct, researchers conducted an extensive literature review followed by a Delphi procedure with HRQoL experts as well as children and adolescent focus groups (Detmar et al., 2006; Ravens-Sieberer, et al., 2005). They determined the level of consensus regarding the conceptualisation and operationalisation of HRQoL (Herdman et al., 2002). They also identified items that measure the relevant construct (HRQoL) and gained direct access to the language and views of participants.

Over and above this, parents also completed questionnaires to determine their perceptions of HRQoL in children and adolescents, which were regarded as complementary to the information from children. Findings from these studies played an important role in the development of the KIDSCREEN-52 questionnaire items which assesses QoL from the

subjective perception of children to ensure that the constructs was adequately presented in the questionnaire. Emphasis was placed on drawing up valid scales, and items were constructed using the language and lay-out preferred by children and adolescents themselves (Ravens-Sieberer et al., 2005). This whole process therefore adds to the validity of the KIDSCREEN-52.

**2.5 Procedures to establish construct validity**

In essence, a test that endeavours to measure HRQoL is attempting to measure a construct (O'Connor, 2004). A construct is a theoretical entity which is not directly observable (Anastasi & Urbina, 1997). It is a hypothetical concept that cannot be measured directly or absolutely confirmed, but it can only be inferred from observations of behaviour (O'Connor, 2004). Cronbach and Meehl (1955) states that construct validation transpires when a researcher believes that his instrument reflects a specific construct, to which are attached certain meanings. The postulated interpretation generates particular testable hypotheses, which are a vehicle of verifying or disconfirming the claim (Cronbach & Meehl, 1955).

However, the acceptability of the construct validity of a measure rests on the acceptability of the construct. Sechrest (2005) maintains that the construct validity of an assessment measure is only as adequate as the construct is acceptable. Hence, one can only have a valid measure of QoL, if the construct has been defined in a convincing, satisfactory way; if the test adequately captures what is implied by the definition; and if the scores on the test are connected to the broader phenomena implicit in the notion QoL (Sechrest, 2005). Thus, the establishment of construct validity must commence with the definition of the theoretical construct (Ghiselli et al., 1981). In addition, Ghiselli et al. (1981) state that the construct is most likely rooted in some theory about behaviour, and the theory elucidates the meaning of

the construct, how it is connected to other constructs, and how it is connected to specific observable behaviours. Construct validity therefore, examines the adequacy of a test in relation to theory (Strauss & Smith, 2009). Accordingly, construct validation can be regarded as a process in which the theory as well as the assessment measure are concurrently evaluated.

Another important component of construct validity is the internal structure of the measure. When dealing with multidimensional measures such as the KIDSCREEN-52, the matter of internal structure is essential to the question of construct validation (Derogatis & Melisaratos, 1983). The analysis of the internal structure of a measure can reveal the extent to which the relationships between test items and test components are consistent with the construct on which the postulated test score interpretations are established (AERA, APA, & NCME, 1999). The degree to which inter-item relationships substantiate the presumptions of the framework is therefore relevant to the validity of the measure (AERA, APA, & NCME, 1999). According to Derogatis and Melisaratos (1983) if a measure is designed to assess a certain number of dimensions defined by means of a specified succession of item sets, then the dimensions defined in this manner should surface in the investigation of a representative empirical problem. The technique usually employed to test the internal structure of a measure is Exploratory Factor Analysis (de Vet, Adèr, Terwee & Pouwer, 2005). This study therefore focused on assessing the construct validity of the internal scaling properties of the KIDSCREEN-52 by means of Exploratory Factor Analysis (EFA).

## 2.6 Conclusion

This chapter provided the basic context for a validation study. It captured the importance of psychometric theory in establishing construct validity evidence of the measure. The next

chapter will explore the literature surrounding the construct HRQoL and the KIDSCREEN-

52 questionnaire.

**CHAPTER THREE**

**LITERATURE REVIEW**

**3.1 Introduction**

This literature review aims to obtain a comprehensive overview of the most pertinent findings in measuring HRQoL as deliberated by other researchers. This chapter focuses on the conceptualisation of HRQoL, HRQoL in children and adolescents, measuring HRQoL within this population, methodological challenges in the assessment of children and adolescent's HRQoL, measurement of HRQoL within a South African context, and application issues relating to the use of assessment measures in a different setting. Lastly, it centres on validation of HRQoL measures and factor analysis as a means of establishing construct validity.
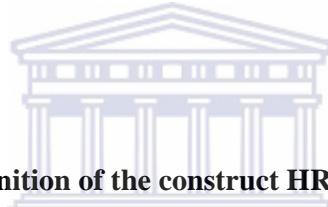
**3.2 Conceptualisation and definition of the construct HRQoL**

The conceptualisation of HRQoL supports a subjective, multidimensional and comprehensive model of health (Ravens-Sieberer et al., 2005). The subjective perception of health and well-being has become a pertinent endpoint in the evaluation of HRQoL. Ravens-Sieberer, Schmidt et al. (2007) state that the subjective information regarding health status is based on the individual's perception of his/her health. According to the WHO Collaborative Cross-National (HBSC) study, how young people feel, is a real and valid aspect of their health, and they can provide an accurate report about it (Currie, Gabhainn & Godeau, 2009). Besides, subjective information regarding health status is not only easily accessible by simply questioning a person, but it is also valuable as it can predict objectively measurable health-related outcomes and as such, embodies a meaningful indicator of health (Ravens-Sieberer as cited in Ravens-Sieberer, Schmidt et al., 2007).

QoL is known to be one of the most complex concepts to define in the social sciences. This can be ascribed firstly to the fact that the concept QoL is often used interchangeably with health and well-being, and life-satisfaction (Okun & Stock, 1987; Tennant et al., 2007; Goldbeck, Schmitz, Besier, Herschbach & Henrich, 2007). Fallowfield (2009) maintains that the construct QoL can be defined in a number of different ways which makes its measurement and integration into scientific study challenging. He further asserts that since illness and its rehabilitation affect the psychological, social and economic well-being, along with the biological integrity of individuals, any definition should be comprehensive while at the same time allowing individual components to be delineated.

Some scholars have equated QoL with health status, well-being (Kaplan et al., 1976), and life satisfaction (Goldbeck et al., 2007) whereas others have defined QoL in terms of the extent to which individuals enjoy the important possibilities in their life (Ranwick, n.d). The World Health Organisation QoL assessment group (WHOQOL, 1995) defined QoL as an individual's perception of his/her position in life in a cultural context, in accordance with the basic social value systems, their objectives, hopes, standards and concerns of life. The construct HRQoL is generally conceptualised as a multi-dimensional psychological concept that encompasses functioning and well-being in the physical, social and psychological or emotional dimensions of life (Arnold et al., 2004; Hays, 2003; Fallowfield, 2009; Rajmil et al., 2004; Ravens-Sieberer et al., 2006; Spieth & Harris, 1996). Health is defined as a state of complete physical, mental, and social well-being and not simply the absence of disease (WHO, 1948). Hays (2003) states that functioning denotes both the ability to perform as well as the performance of everyday routine activities ranging from the most basic self-care activities to more advanced activities such as running a kilometre. Well-being denotes perceptions such as pain and liveliness and how an individual feels about life in terms of

happiness, anxiety, depression, indignation, as well as global perceptions of QoL (Hays, 2003). HRQoL therefore refers to the effect of dysfunction, illness or detrimental circumstances on well-being and life satisfaction relative to an individual's perception of their situation or quandary (Wilkins, O'Callaghan, Najman, Bor & Williams, 2004).

By means of a qualitative analysis, Rajmil et al. (2004) identified the items which are most pertinent in the assessment of HRQoL in children and adolescents. They concluded that all questionnaires analysed included items referring to psychological, social and physical aspects of health. For the purpose of this research HRQoL will be operationally defined as a psychological construct that typifies the physical, mental or emotional, social, psychological and functional aspects of well-being and function from the patient's subjective perspective (Bullinger as cited in Ravens-Sieberer & Bullinger, 1998). This operational definition emphasises the multidimensionality of the construct QoL and the relevance of patients' self-reports (Ravens-Sieberer & Bullinger, 1998).

## 3.3 Health Related Quality of Life in children and adolescents

Young children formulate their own unique conceptions about their health and well-being from a tender age. In order to identify any defects in the health and well-being of children and adolescents, it is therefore necessary to consider how they perceive their own state of affairs. Understanding children and adolescents' perception of health and health issues, in turn, is essential to understanding how they engage with their health, and can bring about short-term as well as long-term population health gains (Keenaghan & Kilroe, 2008).

The current generation of young people encounter many complex challenges that place them at risk and impact on their health and well-being. In a survey conducted by the University of

Nevada Cooperative Extension, the most prominent factors that were perceived as concerns that affect the QoL of young people included teen pregnancy, peer pressure to engage in sexual activities, substance abuse, driving under the influence, parental support and finances to continue education (Singletary, 2005). In addition, they found that gangsterism and harassment from peers (bullying, teasing) were also regarded as significant factors that affect the QoL of young people. Many other studies have also documented factors that impact on young people's QoL such as restricted access to material resources (Ravens-Sieberer et al., 2009), smoking (Rachiotis, Behrakis, Vasilou & Yfantopoulos, 2006), drug and alcohol abuse and related problems (Hubley & Palepu, 2007; O'Hare, 1997), socioeconomic status (Von Reuden, Gosch, Rajmil, Bisegger & Ravens-Sieberer, 2006; Ravens-Sieberer et al., 2009), poor social support (Ravens-Sieberer, Erhart, Gosch & Wille, 2008), and HIV/Aids (Rai, Dutta & Gulati, 2010).
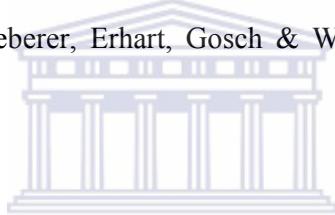
Identifying early childhood determinants of HRQoL allows for the selection of vulnerable children for targeted early intervention projects (Wilkins et al., 2004). The development of HRQoL measures has therefore been encouraged as it serves as a tool to predict and identify those at high risk for health problems, substance abuse, depressive relapse, impaired functioning in school and social relationships (Frisch et al., 2005). In addition, it points out the need for supportive interventions, serves as a prognostic indicator, supports decision-making and informs resource allocation and health care policy (Fallowfield, 2009).

Knowledge regarding the HRQoL of children and adolescents is of particular importance in the public health sector since QoL at this age serves as a springboard for QoL in later life (Bisegger et al., 2005). In this regard, research confirms that childhood problems can manifest in adolescence as well as adulthood. A study conducted by Wilkins et al. (2004)

uncovered early childhood determinants of adolescent HRQoL. Factors such as maternal age, infant attachment patterns, early childhood factors of adversity, and child health and behaviour at the age of five emerged as important determinants of HRQoL during adolescence. Another study by Jirojanakul, Skevington and Hudson (2003) provides evidence that HRQoL is cumulative, and in this sense problems experienced in early childhood and adolescence emerge in adulthood.

Although adolescence has often been described as the healthiest period of life, numerous researchers found that poor health and a high number of symptoms are perceived and reported by a consistent minority of young people (Cavallo et al., 2006; Ravens-Sieberer et al., 2009). Research also indicates gender differences in HRQoL as well as a shift in health status between childhood and adolescence. Cavallo et al. (2006) monitored adolescents' health by means of a multi-national survey involving more than 160 000 participants from 29 European countries, North America and Israel. They found that females are constantly in a worse position than males and older children or adolescents are in a worse position than younger children with respect to their health. These findings are corroborated by findings from other studies. Using the KIDSCREEN-52, Michel et al. (2009) determined age and gender differences in HRQoL among children and adolescents. Results indicate that children on the whole demonstrate better HRQoL values than adolescents, and female adolescents frequently scored lower in HRQoL facets than their male counterparts. Results from other studies by Bisegger et al. (2005) and Ravens-Sieberer et al. (2009) also reveal that in many aspects, children have higher HRQoL than adolescents and with increasing age the HRQoL for females are frequently worse than for males. Bisegger et al. (2005) reported a decrease in HRQoL after the age of 12 in the majority of HRQoL aspects. In the physical and psychological aspects, this decrease was sharper for females than for males.

Decreasing HRQoL during adolescence has been linked by some researchers to developmental issues. Bisegger et al. (2005) ascribe the decline in HRQoL during adolescence to a social and physical transition from childhood to adulthood which often occurs in tandem with transitions in school. Adolescent emotional well-being is affected by biological shifts, cognitive maturation and challenging developmental tasks which may also cause a substantial amount of stress (Goldbeck et al., 2007). Moreover, a full assessment of HRQoL issues in adolescents is problematic due to the variability in the level of maturity within this age group, variation in independence and experience, and probable volatility of emotions (Cramer et al., 1999).

**3.4 Measuring Health Related Quality of Life in children and adolescents**

The importance of HRQoL measurement in health-care research led to the development of numerous generic and disease- or condition specific measures (Rajmil et al., 2004). Disease specific measures assess various aspects of disease and treatment that are pertinent to specific medical conditions (Matza et al., 2004). Disease specific measures are reported to be of greater significance to physicians, focus better on functional areas of particular concern and may hold greater responsiveness to disease specific interventions (Bergner as cited in O'Connor, 1993). However, they are not comprehensive and render the comparison of HRQoL measurements of different diseases difficult (Ravens-Sieberer et al., 2006; Spieth & Harris, 1996).

Generic measures, on the other hand, focus on describing an individual's health independent of the medical condition (Bullinger, Brűt, Erhart & Ravens-Sieberer, 2008), and it is suitable for healthy and clinical populations. Generic measures are designed to be applicable to all population subgroups and allow for comparisons across interventions and diagnostic

conditions, which may include patients who vary in terms of their medical condition (Matza et al., 2004; Davis et al., 2007). Moreover, the absence of measures for a specific disease leaves generic measures as the first choice. However, a recent review of the internationally most widely utilised generic instruments for the assessment of HRQoL in children and adolescents yielded just over 20 instruments (Ravens-Sieberer et al., 2006).

Evidently, even though a myriad of studies have been conducted on HRQoL, the development of assessment measures focused mainly on adults. According to Ravens-Sieberer and Bullinger (1998) only 13% of all HRQoL research publications are related to children and of 320 publications identified, only 9% referred to testing of assessment instruments. However, Spieth and Harris (1996) cite a number of reasons why adult measures are inappropriate for use with children. They contend that the advanced reading levels of adult measures, the necessary level of abstract decision making, the inclusion of dimensions which are irrelevant to children, and the lack of developmental concerns, renders adult measures unsuitable for use with children (Spieth & Harris, 1996). Cramer et al. (1999) caution that merely adjusting the language in a measure designed for adults is unacceptable because adolescents have dissimilar activities and levels of responsibility as well as interests and areas of concerns from adults. According to Bullinger, Schmidt, Petersen and Ravens-Sieberer (2006), QoL dimensions differ with the developmental phases, and whilst the dimensions are generally similar, they tend to differ in their operationalisation across different age stages. The adaptation of adult measures to assess children is therefore questionable (Ravens-Sieberer et al., 2006).

Assessing the HRQoL of children and adolescents must out of necessity take into account developmental issues that impact on QoL. Goldbeck et al. (2007) investigated developmental

aspects of life satisfaction, in particular the effects of age and gender, during adolescence and found a significant decrease in both dimensions of life satisfaction during the course of adolescence. They concluded that decreasing life satisfaction has to be considered as a developmental phenomenon which is a relevant aspect of adolescent well-being and functioning. Moreover, it is necessary to ascertain how children and adolescents' themselves perceive QoL in order to devise appropriate measures. Ideally, the contents of HRQoL instruments must be gained directly from those assessed (Ravens-Sieberer et al., 2005).

However, research exploring the perceptual differences regarding HRQoL among children and adolescents is but few and far between (Zullig, Valois & Drane, 2005). According to Detmar et al. (2006), existing instruments are for the most part founded on expert opinion about what HRQoL comprises and the perceptions of healthy children are rarely included. Zullig et al. (2005) examined adolescent perceptual differences between self-rated health, mental health, physical health and QoL. They found that adolescents rated two distinct constructs and state that self-rated health and QoL should not be utilised interchangeably. They further state that QoL in the context of public high school adolescents' is largely based on self-reported mental health and to a lesser extent on self-reported physical health.

This is precisely why Detmar et al. (2006) conducted focus group discussions in 6 European countries to explore the HRQoL as perceived by children between the ages of 8 and 18. The KIDSCREEN-52 was developed based on findings from this study (Detmar et al., 2006). The contents of the KIDSCREEN-52 were therefore formulated based on the opinions of the envisaged target population, namely, children and adolescents (Ravens-Sieberer et al., 2005; Detmar et al., 2006). In a review of HRQoL measures conducted by Janssens, Gorter, Ketelaar, Kramer and Holtslag (2008) of long-term follow up in children after major trauma,

they identified the KIDSCREEN-52 as one of three suitable measures applicable for a large age range. They confirm that it is reliable and valid, it covers a wide range of the domains of functioning, and the content substantially covers the necessary content as stipulated by the International Classification of Functioning, Disability and Health (ICF) as well as the World Health Organisation (WHO).

## 3.5 Methodological challenges in the assessment of children and adolescent's HRQoL

The assessment of HRQoL has for a long time been neglected and under-investigated in child and adolescent populations (Michel et al., 2009). HRQoL assessment among children and adolescents can therefore be regarded as fairly new. In spite of this, a number of methodological concerns have emerged regarding the assessment of HRQoL in young people. Challenges in the assessment of the HRQoL of children and adolescents include reliability of data, age, literacy level, the use of proxy data and the domains.
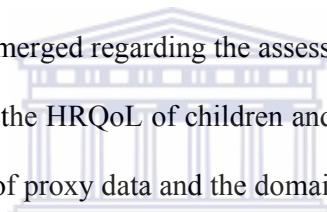
To begin with, doubts have been raised about the reliability of the data provided by children. Claes et al. (2009) states that children are generally regarded as unreliable respondents. This is precisely why earlier endeavours to rate children's HRQoL were derived from data provided by mothers or other proxy reports (Ravens-Sieberer et al., 2005). However, recent research provides evidence that children can reliably and validly self-report their HRQoL when age appropriate measures are used. Varni et al. (2007) examined the feasibility of child-reports and parent proxy-reports of children aged 5-16 years and found that children as young as 5 years old can reliably and validly report their HRQoL when age appropriate measures are used. Another study conducted by Myant and Williams (2005) on children's conceptions of health found that children from as young as four years old are aware of, and are knowledgeable about health and illness, which became more polarised with age. On the

whole, they found that understanding and explanations became more mature and precise with development. This is in line with the Piagetian perspective that children's understanding of health and illness is tied to cognitive development (Papalia, Olds & Feldman, 2004). Another study found children aged 7-8 and 11-12 to be reliable reporters of their health status (Lundqvist et al., 2010). Riley (2004) also found evidence that school aged children can self-report on their health and he states that children's personal reports regarding their own health provide a viable means of monitoring internal experiences of health and distress in childhood as well as adolescence.

The literacy level of children could be an additional impediment in the reliability of their report regarding their health status and HRQoL. Language difficulties can be an obstacle in assessment (Foxcroft & Roodt, 2005). Matza et al. (2004) maintain that the level of language comprehension may place a lower limit on the age appropriateness of a particular questionnaire. Rebok et al. (2001) assessed children's understanding of health-related terms and found that children aged 6-7 had difficulty with some health-related terms, but were able to report on their health experiences. They found that 50% of the five year olds had poor understanding of the key terms presented to them, but by the age of 8 they were able to report on all aspects of their health experiences. Thus, the complexity of the constructs utilised in the test should be age-appropriate as it may play a major role in the validity of the results obtained.

Another methodological issue regarding the measurement of HRQoL revolves around the utilisation of proxy reports. The effectiveness of indirect or proxy reports by parents or health professionals regarding the subjective HRQoL in children and adolescents tends to be a contentious issue. First of all, HRQoL is defined in terms of an individual's subjective

perspective (Ravens-Sieberer et al., 2005) hence using a proxy report is inconsistent with this definition. In reviewing conceptual, methodological and regulatory issues in the assessment of HRQOL, Matza et al. (2004) confirm that HRQoL is subjective and should therefore be assessed from the patient's perspective whenever possible. Others agree that symptoms are by nature subjective and not easy to interpret from the outside or by indirect proxy methods (Lundqvist et al., 2010). Herjanic, Herjanic, Brown and Wheatt (1975) found, by means of a structured interview, 84% agreement between child and parent on questions relating to factual information and only 69% agreement on questions dealing with mental status. The level of agreement between parent and child is therefore dependent on the domain.

In a systematic review of 14 QoL papers, based on 10 different HRQoL measures, Eiser and Morse (2001) substantiate that the level of agreement between parent and child tended to depend on the domain. They found support for the notion that parents are able to judge the child's HRQoL relating to physical rather than emotional or social domains. However, they argue that parents' views could be informed by the burden of care-giving as well as their own concerns and mental health and well-being. They further state that health care professionals will also have limited perceptions of the child's HRQoL, due to the restricted situations in which they are likely to see the child. In a similar review of 19 QoL studies, Upton, Lawford and Eiser (2008) evaluated parent-child agreement for child HRQoL measures and concluded that the differences between parent and child reports could be ascribed to either a lack of parental knowledge regarding children's experiences and beliefs or differences in perception regarding self and others. Even though with cognitively impaired and very young children researchers have turned to proxy-data, no study comparing self- and proxy-completed measures regarded the proxy method to be completely adequate (Fryback, 2003). However, Erhart et al. (2009) maintain that parent proxy reports are favourable with small samples. It

should also be noted that the disparity between parent and child responses to QoL questionnaires should not be regarded to signify a lack of reliability of the concept, but rather as an indicator of two disparate sources of information, which of its own accord is valuable clinical information (Bullinger et al., 2006).

Notwithstanding this, Spieth and Harris (1996) substantiate that self-report measures of HRQoL provide exclusive information that proxy measures from parents or others fails to provide. Externally obtained HRQoL reports or proxy data consequently reflect a doubtful substitute and should only be utilised in exceptional cases or as a supplementary source of information, since they do not accurately embody the manner in which illness is experienced and perceived by children (Ravens-Sieberer et al., 2006).
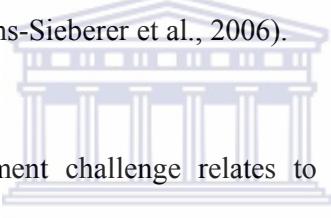
The final measurement assessment challenge relates to the dimensions of HRQoL for children and adolescents. As alluded to previously, QoL dimensions differ with the developmental phases, and whilst the dimensions are generally similar, they tend to differ in their operationalisation across different age stages (Bullinger et al., 2006). There is however agreement that HRQoL is a multidimensional construct which at least contain physiological, psychological and social aspects of well-being (Arnold et al., 2004; Spieth & Harris, 1996; Rajmil et al., 2004; Ravens-Sieberer et al., 2006).

**3.6 The measurement of HRQoL in a South African context**

Similar to the international trend, studies in South Africa centres primarily on adults. The key focus of the majority of South African studies is on a variety of socio-economic and demographic factors that impact on health and well-being.

Møller (2007) conducted a study to review developments that had a bearing on the QoL of ordinary South African's during the transition period following the abolishment of apartheid and concluded that the level of life satisfaction in South Africa coincides with the racial disparities in the standard of living. Similar findings were reported by Westaway (2006), Makiwane and Kwizera (2009) and Møller (2001). Westaway (2006) conducted a study in an informal settlement in Soweto and found that housing was the most important aspect of neighbourhood satisfaction and that respondents' regarded their personal and environmental QoL to be essential for life satisfaction. Another study was done by Møller (2001) in Durban utilising the survey method and in-depth interviews. Results revealed vast differences in living conditions, access to services and life satisfaction across different neighbourhoods. He found that satisfaction with housing and neighbourhood were regarded as major factors contributing to respondents QoL.
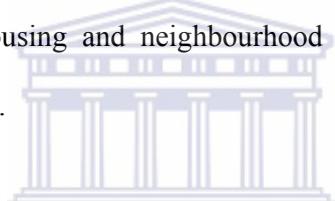
O'Leary (2007) undertook a trend analysis of key subjective and objective variables that have the greatest impact on satisfaction with life for black households. He found that the main factors for life satisfaction were good personal health, good family life, sufficient income, feeling safe and social connectivity. Negative factors that hamper QoL were found to be negative economic conditions, unemployment and crime. Another study also found that crime, unemployment as well as disease, poverty, hunger and HIV/AIDS status depressed QoL in informal settlements (Richards, O'Leary & Mutsonziwa, 2007). The same study established that the most important factors to improve the QoL within these communities were employment, standard of living as well as social connectivity (Richards et al., 2007). Møller (2005) confirms that improved subjective well-being and freedom from fear of crime are significantly correlated. He conducted a victimisation study among 3300 householders in the Eastern Cape to assess the relationship between criminal victimisation and QoL. The

study found that fear of crime and concern about personal safety had a greater negative impact on life satisfaction than actual victimisation. Individual crimes against the person were also found to have a greater negative impact on subjective well-being and beliefs regarding personal safety than property or other household crimes.

Jelsma and Ferguson (2004) focused on the determinants of self-reported HRQoL among culturally and socially diverse South Africans by means of the European Quality of Life – 5 Dimensions (EQ-5D), a generic single index measure validated in Europe as well as Zimbabwe and South Africa. They found that the lowest socio-economic group suffered an excessively high amount of illness and disability. These findings are corroborated by results from another study by Bookwalter and Dalenberg (2004) who found that important differences regarding subjective well-being exist among people based on economic status. A study by Westaway, Olorunju and Rai (2007) sought to determine which QoL domains impact on the happiness of older South Africans and used a structured questionnaire to obtain information on various domains such as financial situations and associations with other people.

An additional factor that impacts on the QoL of South Africans is HIV/AIDS. Due to the stigma attached to the disease, many South Africans are reluctant to present for HIV/AIDS testing. Greef et al. (2010) conducted a longitudinal study to explore the contribution of perceived stigma to the QoL of people living with Aids and found that HIV stigma has a significant negative as well as constant effect on the QoL of individuals living with HIV/AIDS. Another study found a significant relationship among functional abilities, symptom control and QoL among individuals living with HIV/AIDS in Sub-Saharan Africa (Phaladze et al., 2005).

Jelsma, Maclean, Hughes, Tinise and Darder (2005) investigated the impact of highly active antiretroviral therapy (HAART) on the HRQoL of people living with HIV/AIDS (PLWHA) among people in an impoverished suburb in Cape Town, South Africa. The aim of this study was to explore whether HAART is efficacious in improving the self-reported HRQoL in a group of PLWHA by means of the Xhosa version of the EQ-5D. They concluded that, even in a resource-poor environment, HRQoL can be greatly improved by HAART, and that the possible side effects of the drugs seem to have a negligible impact on the well-being of the subjects.

Measures of well-being normed on South African samples include the Minnesota Satisfaction Questionnaire which taps job satisfaction, the Sense of Coherence Scale (for university students and adults) which is used mostly in a research capacity to assess factors which support psycho-social health and well-being, and the Affectometer-2 Scale (AFM2) which is a measure of general happiness and well-being (Foxcroft & Roodt, 2005). Although this measure is applicable to different age, cultural and gender groups, Foxcroft and Roodt (2005) state that it still requires further research within the South African context.

A review of South African literature on QoL and well-being revealed no studies that were conducted on children and only two studies included adolescents as participants. One study assessed the QoL of South Africans with acne vulgaris utilising the Dermatology Specific QoL Questionnaire (DSQL) and respondents' ages ranged from 12 to 47 with a mean age of 23 (Mosam, Vawda, Gordhan, Nkwanyanat & Aboobakr, 2005). They found that South African patients who suffer from acne vulgaris experience significant psychological distress which have an effect on their QoL. The other study examined the level of socio-economic well-being of young South Africans aged 15-34 by measuring domains such as education and

skills, poverty, access to basic services and health care, as well as happiness and satisfaction with life in general (Makiwane & Kwizera, 2009). The findings indicate that the QoL of the majority of young South Africans remain low, and it also reflects the historical racial divides of South African society. The above indicates that a great paucity remains in the literature regarding HRQoL assessment among children and adolescents in South Africa, which spells out the need for the current study.

**3.7 Utilising assessment measures in different settings**

Notwithstanding the prominence HRQoL instruments has gained over the past number of years, Robitail et al. (2006) assert that there is still a dire need for assessment measures that would permit HRQoL assessment at an international level. When a psychometric measure, such as a measure of HRQoL, is utilised in different cultural settings to compare test-takers from diverse cultural backgrounds, we speak of cross-cultural or trans-cultural testing (de Klerk, 2008). As mentioned previously, the lack of appropriate measures to assess the QOL of children and adolescents in South Africa limits researchers to three alternatives: developing a new instrument, adapting and validating an existing instrument or using an existing instrument as is and validating it in the different cultural context (Guillemin, Bombardier & Beaton, 1994; O'Connor, 2004). Economic considerations may dictate which of these options will ultimately be chosen, as it is more expensive to develop a new assessment measure from scratch than to adapt or use an already existing test.

O'Connor (2004) postulates that there is a universal need to evaluate whether measures are being validly applied as test populations differ from the original populations upon which they were developed. He further states that the use of Western measures in non-Western cultures appears particularly problematic since what constitutes QoL is to a large extent influenced by

a person's beliefs and values, and is also largely culturally determined. A cross-cultural qualitative study conducted by Pflug (2009) examined the differences in conceptions regarding the nature of happiness among 57 German and 44 black South African university students by means of a free-format essay on what constitutes happiness to them. Thematic analysis revealed 7 themes, i.e. satisfaction, contentment, positive affect, social relationships, freedom and the opposite of happiness, and an additional theme, 'surprising events', was found with the German respondents. The results indicate that the exact meaning content of these dimensions showed clear influences of individualism/collectivism. In Germany, internal or self-focused factors are more essential for assessing well-being than social or other factors, whereas in a collectivistic culture, such as in South Africa, social harmony and role obligation are key factors in well-being. They also found strong culture-specific factors, such as the impact of philosophical traditions, linguistic influences as well as material living circumstances.
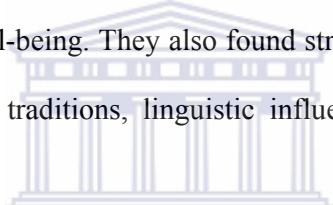
Moreover, cultures might not only vary in the extent of subjective health, but also in the actual complaints expressed and possibly in the exact meaning of the concept (Ravens-Sieberer et al., 2009). According to the South African Employment Equity Act, the utilisation of a psychometric instrument is permissible only when it has been shown to be scientifically valid for respondents from various cultures (Van de Vijver & Rothman, 2004). Thus, the reliability and validity of a measure for a specific context and the suitability of a measure for cross-cultural use are important considerations when using measures (Foxcroft & Roodt, 2005). Furthermore, Foxcroft (2002) emphasises that it is essential for assessment practitioners to be aware of just how unacceptable and unethical it is to use tests without adapting and re-norming them, or to utilise them without establishing 'local' psychometric

properties. There is thus a dire necessity to establish the psychometric properties of the KIDSCREEN-52 within a South African context.

De Klerk (2008) states that measures are adapted from one language and culture to obtain a valid measurement in each culture as well as to facilitate comparative studies across cultural and language groups. Utilising a measure in a different context presupposes that it measures the same construct/s and that the test scores have the same meaning within different populations. However, such comparisons must take place on the same "scale" so as to avoid comparing totally different constructs. According to de Klerk (2008), individual scores based on measures purportedly measuring the same construct in different cultures should not be interpreted at face value. She further suggests that the impact of culture on testing the specific psychological construct ought to be explored in order to adjust measurements to render them meaningful to the particular culture, as well as to obtain comparable or equivalent measures across cultures. According to Van de Vijver and Rothman (2004) a test administered in diverse cultural groups demonstrates structural equivalence if it measures the same underlying construct in all these groups. The statistical technique usually employed to assess structural equivalence is factor analysis. If a measure produces the same factors in diverse cultural groups, there is compelling evidence that the test measures the same construct (Van de Vijver & Rothman, 2004).

The measure of import in this study, i.e. KIDSCREEN-52, is a cross-culturally applicable measure developed as a standardised instrument which can currently be utilised with equal relevance in paediatric and healthy populations (Ravens-Sieberer et al., 2005). It was explicitly tested for the cross-cultural comparability of its measurement (Ravens-Sieberer et al., 2006). To ensure an intercultural comparable assessment, developers of the KIDSCREEN-52 tested whether the interviewees with the same characteristics had the same

probabilities of providing a particular answer to an item, independent of their nationality (Ravens-Sieberer et al., 2005). The results demonstrated that it is possible to assess the HRQoL in a culturally generically comparable manner. However, the cross-cultural comparability was established within different European countries and not in South Africa.

One study conducted outside Europe with the KIDSCREEN-52 was carried out by Hong and associates (2006) in KOREA and cross-culturally validated. They translated the KIDSCREEN-52 following international translation guidelines. They assessed the psychometric properties of the translated test (K-KIDSCREEN-52) by assessing the internal consistency reliability, test-retest reliability, convergent and construct validity. The results indicate that in terms of reliability, Cronbach's alpha for each subscale ranged from 0.77 to 0.95, and test-retest reliability was 0.76. The convergent validity was assessed by comparing the K-KIDSCREEN-52 dimensions with the PedsQL™ 4.0 generic scales and they found high correlations between the physical wellbeing (r=0.32) and moods and emotions (r=0.57, r=0.32) dimensions of the K-KIDCSREEN-52 and similar scales of the PedsQL™ 4.0, but they found low correlations between the school environments (r=0.14) and social support and peers (r=0.22) of the K-KIDSCREEN-52 and similar scales of the PedsQL™ 4.0, which warrants further explanation. They concluded that the measure was suitable for use in Korea as the correlation coefficients between the K-KIDSCREEN-52 and PedsQL™ 4.0 was high for the assessment of similar dimensions.

### 3.8 Validation of HRQoL measures

According to Spieth and Harris (1996), the utilisation of psychometrically sound QoL measures is a significant prerequisite for achieving valid QoL outcomes. Hence, apart from the development of well-validated measures, the demonstration of the psychometric

properties of existing measures is particularly encouraged. Anastasi and Urbina (1997) highlight a number of different ways of establishing construct validity which include correlations with other tests, factor analysis, internal consistency and convergent and discriminant validation. Validity should be established with reference to the specific use for which the test is being considered and the choice of validation procedure is governed by the use to be made of test scores (Anastasi & Urbina, 1997).

A number of previous research on construct validity of well-being measures demonstrated construct validity through a multitrait-multimethod approach (Lix, Metge & Leslie, 2009); convergent and discriminant validity (Burkhardt, Anderson, Archenholts & Häg, 2003; Torres et al., 2009); correlations between a new test and similar earlier tests (Duarte, Ciconelli & Sesso, 2005; Hong et al., 2006; Korevaar et al, 2002); confirmatory factor analysis (CFA) (Abbott et al., 2006; Erhart, Ellert et al., 2009; Stewart-Brown et al., 2009) and structural equation modelling, which includes factor analysis (Fayers & Hand, 1997). Other researchers mostly used EFA (de Vet et al., 2005). EFA was found to be the most common form of factor analysis in educational as well as health status assessment (de Vet et al., 2005; Pohlmann, 2004).

A study conducted by Tennant, Hiller, Fishwick and Platt (2007) used confirmatory factor analysis to test the hypothesis that the Warwick-Edinburgh Mental Well-being Scale (WEMWBS) tests a single construct. They assessed the internal consistency of the WEMWBS using Cronbach's alpha and content validity by reviewing the frequency of complete responses as well as the distribution of responses to each item. They also assessed the criterion validity through correlations between the WEMWBS and other scales by testing whether the scale discriminated between population groups in line with pre-specified

hypotheses. In another study, Stewart-Brown et al. (2009) assessed the internal construct validity of the WEMWBS by means of Rasch analysis.

Ravens-Sieberer, Auquier et al. (2007) examined the construct validity of the KIDSCREEN-27 by testing a priori presumed associations with other generic HRQoL measures (i.e. YQOL-S, PedsQL and CHIP), socio-economic status and indicators of mental and physical health. They also assessed the criterion validity of the KIDSCREEN-27 by examining the correlation with and the percentage explained variance with corresponding scales of the KIDSCREEN-52 instrument. Correlations between the KIDSCREEN-27 dimensions and the other HRQoL measures were moderate to high ($r = 0.36$ to $0.63$) for those assessing similar constructs.
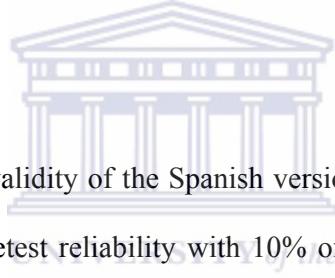
To examine the reliability and validity of the Spanish version of the KIDSCREEN-52, Tebe et al. (2008) assessed the test-retest reliability with 10% of the sample 2-4 weeks after the initial administration. Cronbach's alpha values ranged from 0.74 to 0.86 and the intraclass correlation coefficients varied from 0.55 to 0.79; Rasch analysis indicated a good item-fit in each dimension, and CFA confirmed the underlying structure of the 10 dimensions model. They concluded that the measure had satisfactory levels of reliability as well as validity for the Spanish population.

Erhart, Ravens-Sieberer, O'Dickinson and Colver (2009) assessed the Rasch measurement properties and the differential item functioning (DIF) of the KIDSCREEN-52 in children with cerebral palsy. Existing data from the KIDSCREEN project (3219 children and 2126 parents) and the SPARCLE (Study of Participation of Children with Cerebral palsy living in Europe) project (501 children and 823 parents) were used. All the items of the KIDSCREEN-52 were

found to fit the partial credit model (smallest *p*-value 0.256). Only one item of the KIDSCREEN –proxy parent version did not fit the data well (smallest *p*-value 0.001) and statistically significant DIF was detected in some items, but the practical impact of DIF was small.

**3.9 Exploratory factor analysis as a means of construct validation of QoL measures**

An important stage in the validation of multi-item measures such as the KIDSCREEN-52 is factor analysis. Factor analysis is a procedure which is designed to uncover whether or not the pattern or responses on a number of items can be explained by a smaller number of underlying factors. Exploratory Factor Analysis (EFA) continues to be one of the standard and most extensively utilised methods for demonstrating construct validity (Fayers & Hand, 1997). Numerous studies in the field of QOL employed EFA analysis to establish construct validity (Della, De Joy, Goetzel, Ozminkowski & Wilson, 2008; Gentile et al., 2008; Panepinto, Hoffman & Pajewski, 2009; Sapin, Simeoni, Khammar, Antoniotti & Auquier, 2005; Springer, McQueen, Quintanilla, Arrivillaga & Ross, 2009; Yi et al., 2008).

Sapin and colleagues (2005) assessed the construct validity of the VSP-A,a HRQoL instrument  by utilising inter-item correlations, item-dimension correlations, and principal component analysis with Varimax rotation. Della et al. (2008) conducted EFA using the principle components method of factor estimation with oblique (Oblimin) factor rotations. They state that they utilised the principle components method since it is less prone to suffer from factor indeterminacy, and oblique rotations allowed factors to covary. In another study, Skevington, Lofty and O'Connell (2004) assessed the construct validity of the WHOQOL-BREF by first carrying out an EFA (with Varimax rotation) to explore the factor structure of

the data. CFA were then conducted, using structural equation modelling, to obtain objective measures of model fit.

In a critical assessment on the use of factor analysis in the field of health sciences, de Vet and colleagues (2005) reviewed 13 different journals to appraise a well-known health and well-being questionnaire, the Short Form 36 health survey questionnaire (SF-36). Of the 28 studies identified, 22 studies employed EFA and 5 CFA whilst one used both. However, they contend that in 15 of the 23 studies CFA would have been more appropriate and they developed a checklist for the appropriate use of either EFA or CFA. In spite of this, they regard EFA appropriate if the aim of the study is to examine the factor structure of a health status questionnaire in a population or language in which the measure had not yet been used without a prior hypothesis. Accordingly, since the current study aims to ascertain whether the items of the KIDSCREEN-52 can be categorised into factors signifying the different dimensions of the construct HRQoL, EFA can be regarded as an appropriate statistical procedure to reveal the underlying structure of this instrument.

It should, however, be noted that some researchers either caution the use of factor analysis (de Vet et al., 2005) whilst others openly express their disapproval and regard it as an inappropriate method to establish construct validity (Kaplan et al., 1976). In support of factor analysis though, Claes et al. (2009) endorse the fact that a QOL assessment instrument should be validated by demonstrating the factor structure of the model's domains and ascertaining the statistical characteristics of the domains and measurement indicators. Moreover, Goodwin (2000) postulates that FA is commonly utilised to demonstrate construct validity, as a mechanism for the empirical development of theory concerning the nature of constructs, and

to summarise groups of highly correlated variables into smaller factors in order to facilitate subsequent analysis.

As regards EFA and the KIDSCREEN measures, Robitail and associates (2007) assessed the validity of the KIDSCREEN-27, which was administered in 13 European countries, by testing its multi-dimensional structure, the unidimensionality of its five dimensions and its internal consistency. They performed EFA, using principal component analysis (PCA), on the sample as a whole. For the combined sample from all the countries, EFA with procrustean rotations revealed a five-factor structure which explained 56.9% of the variance. CFA revealed an acceptable model fit (RMSEA = 0.068, CFI = 0.960) and the unidimensionality of all items was confirmed (INFIT: 0.81-1.15). Another study examined the psychometric properties of the KIDSCREEN-10 index, which was developed from the longer KIDSCREEN-52 and -27 (Erhart, Ottova et al., 2009). They wanted to ascertain whether the psychometric properties would still be retained if the items of the KIDSCREEN-10 were applied alone and not within the context of the longer KIDSCREEN-52. Results indicate that the KIDSCREEN-10 items functioned as good indicators irrespective of the context in which they were applied, but it also revealed noticeable differences between the 15 countries in which the studies were conducted. The dimensionality of a measure is commonly assessed through factor analysis, which provides insight into the interrelationships among variables and the underlying structure of the data, and is a good starting point for many other multivariate techniques (Hair, Black, Babin & Anderson, 2010). Erhart, Ottova et al. (2009) concluded that a limitation of this study concerns the fact that no explicit test of unidimensionality and differential item functioning of the KIDSCREEN items across countries were performed.

**3.10 Conclusion**

The current chapter looked at the available literature on HRQoL and issues relating to measuring HRQoL. The following chapter will focus on the methodology of the primary study as well how the current study was conducted.

## CHAPTER 4

## METHODOLOGY

**4.1 Introduction**

The current study forms part of a larger study outlined in chapter one, viz. the HECVW study. This construct validation study therefore employed Secondary Data (SD) in order to examine the construct validity of the KIDSCREEN-52, one of the measures employed in the larger study. In addition, it made use of sophisticated data analytic methods, which rendered it quantitative in nature. The quantitative approach focuses on reliability (Welman, Kruger & Mitchell, 2005) and producing generalisable results (Marshall, 1996). Quantitative research is very structured and controlled which allows the researcher to identify and isolate variables (Welman et al., 2005) which is fitting for the current study as assessing construct validity by means of factor analysis entailed condensing a group of empirical indicators into a smaller set of latent variables (Claes et al., 2009).
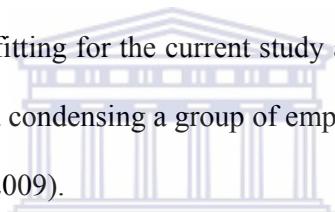
Even though the use of existing data is recommended whenever such data provides appropriate information, Judd, Smith and Kidder (1991) advise that all the major issues that are of concern in the collection of original data (e.g. adequacy of the sample) must also be assessed when using existing data. Accordingly, the procedures and methods of the main study as well as the method of investigation employed in the current study, is outlined in this chapter.

**4.2 Research design**

The researcher conducted a Secondary Data Analysis (SDA)**,** which can be described as the analysis of data collected previously by another researcher (Judd et al., 1991). Fife-Shaw (2006) defines SDA as the re-analysis of original data, collected earlier by another research

study, utilising different statistical procedures in order to examine alternative questions. Secondary analysis allowed the researcher of the current study to pursue a particular research interest, i.e. establishing the construct validity of the KIDSCREEN-52 by analyzing the existing survey data collected by the broader study conducted by Savahl et al. (2009). This research can therefore be deemed to commence where the previous one ended. Moreover, contemporary research in all the social sciences largely involves the analysis of data that were previously gathered by other investigators (Judd et al., 1991). This can primarily be ascribed to the many benefits that may be derived from doing SDA.

In line with the benefits delineated by Babbie and Mouton (2001), the researcher of the present study avoided the enormous expenditure of time and money which such surveys entail. Furthermore, unnatural reactions in the behaviour of respondents to the fact that they are being observed (i.e. reactivity) are avoided in secondary analysis (Singleton et al., 1993). Also, whilst all knowledge builds upon the work of others who have gone before, SDA allows this to happen in a much more direct and practical sense than usual (Dale, Arber & Procter, 1988) especially in the case of the current study which built on the main study. Stewart (1984) holds that researchers using SDA develop a sound scientific scepticism with regards to data collected by other researchers. In addition, Mouton (2001) states that SDA compels the researcher to be explicit about the assumptions and theory that underlie the data; as is the case in the current construct validation study which is framed by psychometric theory.

### 4.3 Sampling

Data for this construct validation study was drawn from the HECVW study. The primary study employed stratified interval criterion sampling. Stratified random sampling enabled the

researchers to divide the areas into three strata (high, medium and low-risk violence areas) which were then further divided (into high-and- low income areas) based on South African Police Services (SAPS) statistics. Schools were purposively sampled from these strata. From the sub-samples, schools were then randomly selected from a list by selecting every third school. The criteria used in the choice of sample was that participants should be grade nine learners from six public schools within the Education Management and Development Centre (EMDC) South Metropole of the Western Cape Education Department.

## 4.4 Participants

The 565 participants of the main study comprised 348 female and 218 male grade 9 learners aged 14-18. Since the current study examined the construct validity of the KIDSCREEN-52 by means of EFA, the researcher opted to retain the full sample of the broader study. An important element which may impact the stability of factor analysis results, is sample size, in that the larger the sample the more stable the results obtained from EFA (Hammond, 2006). To provide a reliable factor solution, Hair et al. (2010) advise that a sampling size of at least 100 or larger should be used and as a general rule there should be at least five times as many subjects as variables. However, they maintain that a more acceptable size would have a 10:1 ratio. Since the data set for the current study comprised 565 participants the sample size was suitable to conduct EFA in relation to the number of variables (52) in the KIDSCREEN-52. Tables 1 to 4 below present the sample in terms of age, gender, language and area.

**Table 1:**

**Distribution of participants per age**

| Age | | Frequency (n) | Percent | Cumulative percentage |
|---|---|---|---|---|
| **Valid** | 14 | 147 | 26.0 | 26.0 |
| | 15 | 314 | 55.6 | 81.6 |
| | 16 | 62 | 11.0 | 92.6 |
| | 17 | 27 | 4.8 | 97.3 |
| | 18 | 15 | 2.7 | 100.0 |
| **Total** | | 565 | 100.0 | |

Table 1 above disaggregates the sample in terms age. The ages range from 14 to 18 years old. This table indicates that the sample of the present study consisted mainly of 14 (26%) and 15 (55.6%) year olds.
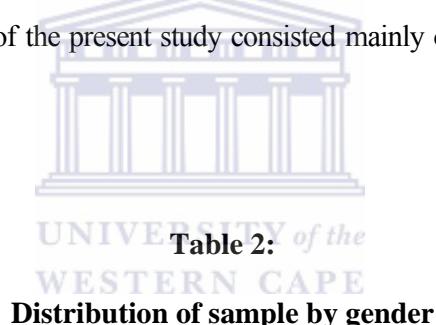
**Table 2:**

**Distribution of sample by gender**

| Gender | | Frequency (n) | Percentage | Valid Percent | Cumulative percentage |
|---|---|---|---|---|---|
| **Valid** | Male | 215 | 38.1 | 38.2 | 38.2 |
| | Female | 348 | 61.6 | 61.8 | 100.0 |
| | Total | 563 | 99.6 | 100.0 | |
| **Missing** | System | 2 | .4 | | |
| | **Total** | 565 | 100.0 | | |

Table 2 above presents the sample disaggregated by gender. This table indicates that the sample consisted of 215 (38.1%) males and 348 (61.6%) females. In addition, two participants did not indicate their gender.

**Table 3:**

**Distribution of sample by language**

| | Language | Frequency | Percent | Cumulative Percent |
|---|---|---|---|---|
| | English | 299 | 52.9 | 54.2 |
| | Afrikaans | 78 | 13.8 | 68.3 |
| **Language Group** | Xhosa | 144 | 25.5 | 94.4 |
| | Other | 6 | 1.1 | 95.5 |
| | English/Afrikaans | 17 | 3.0 | 98.6 |
| | English/Xhosa | 8 | 1.4 | 100.0 |
| | Total | 552 | 97.7 | |
| **Missing** | System | 13 | 2.3 | |
| **Total** | | 565 | 100 | |

Table 3 above displays the composition of the sample in terms of language. English was the predominant language of more than half of the participants. However, a significant percentage of students spoke Xhosa (25.5%) and Afrikaans (13.8%) which together constitutes more than 39% of the sample. Since the KIDSCREEN-52 was administered in English, without any adaptations in terms of language and cultural concerns, this could play a significant role in the results obtained from the current study.

**Table 4:**

**Distribution of sample by area in terms of risk and exposure to violence**

| | | Frequency | Percent | Cumulative Percent |
|---|---|---|---|---|
| **Valid** | 1 | 122 | 21.6 | 21.6 |
| | 2 | 140 | 24.8 | 46.4 |
| | 3 | 303 | 53.6 | 100.0 |
| **Total** | | 565 | 100.0 | |

Table 4 above disaggregates the sample into those from high (3), medium (2) and low (1) risk areas based on SAPS police statistics in terms of risk and exposure to violence. It is evident from the table above that the majority of participants (53.6%) come from high risk areas.

## 4.5 Procedure

Ethical clearance was obtained for the main study from the University of the Western Cape Research Ethics Committee, as well as consent from the Western Cape Education Department and the respective principals so as to gain access to the schools. Participants and their guardians were informed about the aims and objectives of the study, given an information sheet and invited to participate in the study without coercion. Thereafter, parents' signed consent (Appendix A) and learners' assent (Appendix B), to participate in the study, was obtained and only learners whose parents completed the forms were allowed to participate in the study. The school based surveys were administered by the research team in the presence of staff members of the respective schools. The test administrators gave students clear, standardised instructions regarding completion of the measure. The test battery took about 45 minutes to complete. The data were stored in a secure place and was only accessible to the research team.

## 4.6 Data Collection Instrument (KIDSCREEN-52)

The measure of import in the current study is the KIDSCREEN-52 (Appendix C). This self-report measure assesses children and adolescents' subjective health and well-being and is applicable to healthy and paediatric (ailing or disabled) populations from 8 to 18 years of age.

The questionnaire assesses the frequency of behaviour/feelings or, the intensity of an attitude by using a 5-point Likert response scale, and the recall period is one week. This instrument

can be used for different paediatric populations as well as healthy children and adolescents. The aim is to identify children and adolescents who are at risk regarding their subjective health and to present appropriate early interventions by integrating the measure in health services research and health reporting (Ravens-Sieberer et al., 2005). This measure assesses 10 dimensions, namely physical well-being, psychological well-being, moods and emotions, self-perception, autonomy, parent relations and home life, social support and peers, school environment, social acceptance (bullying) and financial resources.

Aside from the full version (52 items with 10 dimensions), there is also an abbreviated version (27 items with 5 dimensions), as well as a 10-item index of general HRQoL (Ravens-Sieberer et al., 2006).

### 4.6.1 Dimensional Structure of the KIDSCREEN-52

According to Anastasi and Urbina (1997) almost any information collected in the course of developing or using a measure is relevant to its validity. If the items and instrument content adequately represent the domain they are expected to measure, if test construction procedures indicates that all dimensions generally regarded relevant have aided in defining the domain/s and that the domain/s was appropriately sampled, then this serves as content validity evidence (Kaplan et al., 1976). Also, if the chosen theoretical construct for a health status index directs the choice of the content, and the content buttresses the construct, it also serves as content validity evidence (Kaplan et al., 1976). As mentioned previously, in developing the KIDSCREEN-52 questionnaire, researchers first conducted a literature review and then a Delphi panel of experts was convened and they reached consensus regarding the conceptualisation and operationalisation of HRQoL (Herdman et al., 2002; Ravens-Sieberer et al., 2006). They identified physical, psychological and social aspects of health as the broad

domains to be retained in the assessment of QoL. It was also agreed that the HRQoL assessment measure had to cover a multidimensional concept, reflecting the respondents' views of their state of health.

In addition, Detmar et al. (2006) explored the perceptions of children to ascertain their perceptions about what constitutes HRQoL. Results indicates that social functioning and social relations (i.e. social dimension), are the most significant aspects of HRQoL for children and adolescents. The physical dimension becomes important when limitations on activity or handicaps are present, whilst social rejection and bullying also emerged as an important topic in a general HRQoL questionnaire. Thus, the design of the KIDSCREEN-52 took into account the specific aspects of children and adolescent's lives that comprise the important dimensions to be measured as well as the relevant items that correspond to the target population's experiences, activities and relevant context for assessing their HRQoL.
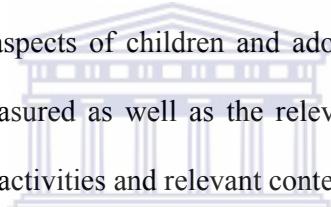
The KIDSCREEN-52 is conceived of as measuring the following ten different HRQoL dimensions (Detmar et al., 2006; Ravens-Sieberer et al., 2005):

Dimension 1: *Physical well-being*

This dimension investigates the level of the respondent's physical activity, energy and fitness. The level of physical activity is measured with regards to the child or adolescent's capacity to get around the school and home, and to play or do physically challenging activities such as sports, as impairment also impacts on physical activity. It also assesses the capacity for lively or energetic play; and looks at the extent to which individuals feel unwell and complain about poor health.

Dimension 2: *Psychological well-being*

This dimension explores the psychological well-being of the child/adolescent, including positive emotions and satisfaction with life. In particular, it reveals the positive emotions and perceptions experienced by the individual. The questions examine the extent to which individuals' experience positive feelings such as joy, happiness and cheerfulness; and their view of their satisfaction with life up till now.

Dimension 3: *Moods and emotions*

This dimension covers how often the child/adolescent experiences depressive moods and emotions, and stressful feelings. It particularly exposes feelings such as sadness, resignation, loneliness and adequacy/inadequacy as well as the extent of such feelings. This dimension indicates a high score for QoL if these negative emotions are rare.
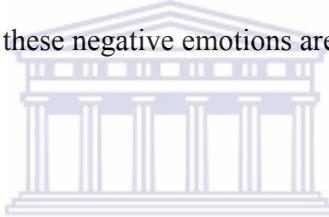
Dimension 4: *Self-perception*

This dimension taps the child or adolescent's perception of self. It explores the extent to which respondents view their bodily appearance positively or negatively and their level of satisfaction with their looks or appearance as well as clothing and accessories. It is meant to reflect the value individuals assign to themselves and their opinions regarding how positively they regard others to value them.

Dimension 5: *Autonomy*

This dimension looks at the opportunities presented to a child/adolescent to create social and leisure time and whether they feel sufficiently provided with such opportunities. It examines their level of autonomy, which is regarded as a vital developmental issue in creating an individual's identity. Autonomy refers to the individual's freedom of choice, independence

and self-sufficiency. It particularly examines the degree to which the child or adolescent feels they are able to shape their lives and make decisions regarding day-to-day activities.

Dimension 6: *Parent relations and home life*

This dimension examines relationships with parents and the atmosphere at home. It explores the value of interaction between child/adolescent and parent/caregiver, and their feelings towards parents or caregivers. Emphasis is placed on the extent to which the child/adolescent feels loved and supported by the family, whether the home environment is comfortable or not and whether they think they are treated fairly.

Dimension 7: *Social support and peers*

This dimension looks at the nature of the respondents' relationships with other children or adolescents. Social relationships with peers and friends, quality of the interaction, perceived support and acceptance from friends and peers, and their ability to form and create friendships, are examined. Aspects regarding communication with others are looked at as well as the extent to which the child/adolescent experiences positive group feelings. The extent to which they feel part of a group and respected by friends and peers, are also examined.

Dimension 8: *School environment*

This dimension explores the child's/adolescent's views of their cognitive ability, learning and concentration, and their opinion about school. It encompasses the child/adolescent's satisfaction with his/her ability and performance at school, general feelings about school as well as opinions about relationships with teachers.

Dimension 9: *Social acceptance (bullying)*

This dimension entails the aspect of feeling rejected by peers in school. It explores the feeling of being rejected and of anxiety towards peers.

Dimension 10: *Financial resources*

This dimension examines the child/adolescent's perception of their financial resources. It explores whether the individual feels s/he has sufficient financial resources to enable him/her to have a lifestyle similar to other children/adolescents and present the opportunity to do things with peers.

Each dimension consists of specific items as outlined in table 5 below.

**Table 5:**

**Dimensions and items of the KIDSCREEN-52**

| Dimension | Items |
|---|---|
| 1. Physical well-being | 1. In general, how would you say your health is?<br>2. Have you felt fit and well?<br>3. Have you been physically active?<br>4. Have you been able to run well?<br>5. Have you felt full of energy? |
| 2. Psychological well-being | 1. Has your life been enjoyable?<br>2. Have you felt pleased that you are alive?<br>3. Have you felt satisfied with your life?<br>4. Have you been in a good mood?<br>5. Have you felt cheerful?<br>6. Have you had fun? |
| 3. Moods and emotions | 1. Have you felt that you do everything badly?<br>2. Have you felt sad? |

| | |
|---|---|
| | 3. Have you felt so bad that you didn't want to do anything? |
| | 4. Have you felt that everything in your life goes wrong? |
| | 5. Have you felt fed up? |
| | 6. Have you felt lonely? |
| | 7. Have you felt under pressure? |
| 4. Self-perception | 1. Have you been happy with the way you are? |
| | 2. Have you been happy with your clothes? |
| | 3. Have you been worried about the way you look? |
| | 4. Have you felt jealous of the way other girls and boys look? |
| | 5. Would you like to change something about your body? |
| 5. Autonomy | 1. Have you had enough time for yourself? |
| | 2. Have you been able to do things that you want to do in your free time? |
| | 3. Have you had enough opportunity to be outside? |
| | 4. Have you had enough time to meet friends? |
| | 5. Have you been able to choose what to do in your free time? |
| 6. Parent relations and home life | 1. Have your parent(s) understood you? |
| | 2. Have you felt loved by your parent(s)? |
| | 3. Have you been happy at home? |
| | 4. Have your parent(s) had enough time for you? |
| | 5. Have your parent(s) treated you fairly? |
| | 6. Have you been able to talk to your parent(s) when you wanted to? |
| 7. Social support and peers | 1. Have you spent time with your friends? |
| | 2. Have you done things with other girls and boys? |
| | 3. Have you had fun with your friends? |
| | 4. Have you and your friends helped each other? |
| | 5. Have you been able to talk about everything with |

| | |
|---|---|
| | your friends? |
| | 6. Have you been able to rely on your friends? |
| 8. School environment | 1. Have you been happy at school? |
| | 2. Have you got on well at school? |
| | 3. Have you been satisfied with your teachers? |
| | 4. Have you been able to pay attention? |
| | 5. Have you enjoyed going to school? |
| | 6. Have you got along well with your teachers? |
| 9. Bullying | 1. Have you been afraid of other girls and boys? |
| | 2. Have other girls and boys made fun of you? |
| | 3. Have other girls and boys bullied you? |
| 10. Financial Resources | 1. Have you had enough money to do the same things as your friends? |
| | 2. Have you had enough money for your expenses? |
| | 3. Do you have enough money to do things with your friends? |
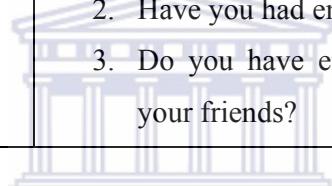
## 4.6.2 Psychometric properties of the KIDSCREEN-52

There are two technical requirements that any measuring instrument must meet, i.e. reliability and validity. Reliability refers to the consistency with which an instrument measures whatever it measures, or yields the same results when applied repeatedly to the same object (Babbie & Mouton, 2009; Foxcroft & Roodt, 2005). Validity refers to the extent to which a scale or set of measures actually measures the characteristics or dimension we intend to measure (Walsh & Betz, 2001).

Multinational studies conducted by Ravens-Sieberer et al. (2005) have demonstrated that the KIDSCREEN-52 has high reliability and validity. To assess the internal consistency reliability of the KIDSCREEN-52, they calculated Cronbach's alpha coefficients for the ten

KIDSCREEN dimensions which ranged satisfactorily from 0.77 to 0.89 (Ravens-Sieberer et al., 2005). Only for one scale (social acceptance/bullying), the Cronbach's alpha was found to be below 0.70 in one country (France).

The convergent validity (i.e. measures that should be related are in reality related) was assessed by comparing the KIDSCREEN-52 dimensions and the KINDL-R scales (Ravens-Sieberer et al., 2005). Convergent validity was regarded as established when correlations between comparable dimensions were considerably higher than between theoretically dissimilar dimensions, and were of reasonable magnitude. Results indicated high correlations for all similar concepts or dimensions such as the KIDSCREEN-52 physical well-being dimension and physical functioning scale of the KINDL-R ($r = 0.53$), which indicates a satisfying convergent validity. Theoretically expected low correlations (i.e. divergent validity) were in fact found for the KIDSCREEN-52 financial resources dimension with the KINDL-R scales. In addition, social acceptance (bullying) demonstrated low coefficients in nearly all analyses except with the Friend's scale of the KINDL-R (Ravens-Sieberer et al., 2005). In another study the KIDSCREEN-52 showed moderate correlations with the subscales of the YQOL-S (Youth Quality of Life Instrument – surveillance version) and the CHIP-AE (Child Health and Illness Profile – Adolescent Edition (Ravens-Sieberer, Schmidt et al., 2007).

In another validation study on the KIDSCREEN-27, the abbreviated version of the KIDSCREEN-52, Robitail et al. (2007) examined the multi-dimensional structure, internal consistency reliability and the unidimensionality of its five dimensions. However, since no psychometric properties have been established within a South African context to date, items and dimensions of the KIDSCREEN-52 remain relevant to European children and

adolescents from participating countries only (Ravens-Sieberer et al., 2005). This, again, points to the dire need for the current study.

## 4.7 Data analysis

### 4.7.1 Introduction

In view of the fact that this proposed study employed SDA, the researcher utilised the existing data of the HECVW study to conduct different statistical tests using the Statistical Program for the Social Sciences (SPSS, version 17.0) package. To examine the construct validity of the KIDSCREEN-52 within a South African context, this research explored its factor structure, and the internal consistency reliability of the subscales. Since an accumulation of missing values due to paired missing values may critically curtail the number of subjects on which the variance co-variance matrix is based (de Vet et al., 2005), the researcher will first provide information about the number of missing values.
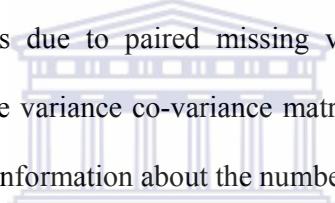
### 4.7.2   Examining the factor structure of the KIDSCREEN-52

**Research Aim 1: To examine the factor structure of the KIDSCREEN-52 within a South African context to determine whether this study replicates the 10-dimensional structure identified by previous European studies.**

The method used for examining the factor structure of the KIDSCREEN-52 was the statistical technique of exploratory factor analysis (EFA). The motivation for using EFA is to determine whether this study replicates the 10-dimensional structure identified by previous European studies.

**4.7.2.1 Factor analysis**

Factor Analysis is one of the most widely used statistical techniques for exploring the underlying structure of a set of variables. It is a multivariate statistical technique that is utilised to investigate the relationship among variables. Hair et al. (2010) explain that factor analysis is an interdependence technique with the primary objective of defining the underlying structure of the variables within the analysis. It supplies the mechanisms for investigating the structure of the interrelationships (correlations) between large numbers of variables (e.g. questionnaire responses) by defining groups of variables that are highly correlated, known as factors (Hair et al., 2010). These factors are assumed to represent dimensions within the data. Factor analysis is therefore appropriate if the aim is to determine the factor structure (dimensions) tapped by a questionnaire or to examine if the questionnaire demonstrates the same dimensions across different groups (de Vet et al., 2005).
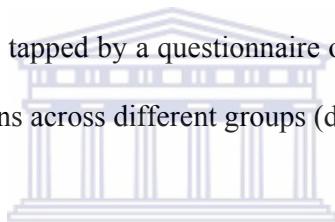
Factor analysis therefore identifies the clusters of high correlation coefficients between subsets of a measure. It establishes how much variance they have in common and the extent to which each item contributes to this common variance. Thus, a large group of items can be reduced to a smaller set of factors, which can elucidate the highest amount of common variance in the bivariate correlations between them (Blaikie, 2003). Factor analysis can therefore ascertain whether all the items are highly correlated and can be viewed as constituting a common factor or a single scale (Blaikie, 2003). It can also establish whether subsets of items constitute separate scales, or whether subscales are present in a general scale. Given the fact that construct validity is anchored in the logical relationships among variables (Babbie, 2004); factor analysis is the ideal statistical method for the current study because it is used to analyse the interrelationship among variables.

There are basically two broad types of factor analysis: exploratory and confirmatory factor analysis. Confirmatory factor analysis (CFA) is regarded as more appropriate when prior hypotheses exist, founded on theory or earlier analysis (de Vet et al., 2005). EFA, on the other hand, is utilised when the objective is pure data reduction or the examination of the factor structure (dimensions) being assessed by a questionnaire (de Vet et al., 2005). Thus, the statistical analysis employed to examine the underlying structure of the KIDSCREEN-52, and to determine whether we find the same number of subscales identified by previous European studies was EFA.

**4.7.2.2 Design issues and assumptions in factor analysis**

In factor analysis one has to consider several design issues in order to obtain successful results, namely the sample size required, variable selection and correlations among variables. Since factor analysis is a variance partitioning technique, an adequate sample size is required to minimise sampling error. To produce a reliable factor solution, it is advisable that a sample size of 200 or larger is used where possible (Hammond, 2006). As a general rule, it is also recommended that there are at least five times as many observations as the number of variables to be analysed, and the more acceptable sample size would have a ratio of 10:1 whereas the minimum absolute sample size should be 50 observations (Hair et al., 2010). As mentioned previously, our sample size is slightly more than a ratio of 10:1 (565 ÷ 52) observations per variable which falls within acceptable limits and provides an adequate basis for the calculations between variables. According to Hair et al. (2010) the higher number of cases-per-variable ratio minimise the chance of overfitting the data, i.e. deriving factors that are sample specific with little generalisability.

The next design issue pertains to variable selection. The important consideration here is the

type of variables used and the number of variables included in the analysis. All the variables in the data of this study are item level data on a Likert scale. Hair et al. (2010) suggest that one should ensure that you include several variables (five or more) that may represent each factor. The current study included 52 variables with eight of the factors represented by five or more variables and two factors represented by three variables.

The next design issue centres on the intercorrelations among variables. According to Field (2005), to run a factor analysis, one needs to have variables that correlate reasonably well, but not perfectly. Thus, before the researcher ran the factor analysis, she first determined the factoriability of the data. This was achieved by looking at the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO), the Bartlett's Test of Sphericity and the diagonal elements of the anti-image correlation matrix (Field, 2005; Hair et al., 2010). Measured by the Kaiser-Meyer-Olkin (KMO) statistics, sampling adequacy predicts if data are likely to factor well, based on correlation and partial correlation. KMO varies from 0 to 1.0 and Kaizer recommends that a bare minimum KMO overall should be .50 or higher to proceed with factor analysis (Field, 2005). According to Field (2005) values between .70 and .80 are regarded as good and between .80 and .90 are deemed great, whilst values above .90 are superb. The Bartlett's Test of Sphericity tests whether the original correlation matrix is an identity matrix. For FA to work we need some relationship between variables and if the *R*-matrix were an identity matrix then all correlation coefficients would be zero. It is also important to examine the diagonal elements of the anti-image correlation matrix. The value should be above the bare minimum of .50 for all variables (and preferably higher) (Field, 2005). Thereafter, one needs to look at the correlation coefficients themselves to check for values greater than .9 which is indicative of singularity in the data.

There are also a number of underlying statistical assumptions that impact FA to the extent that they affect the derived correlations. These assumptions that need to be examined are linearity, lack of high multi-collinearity (although some degree of multi-collinearity is desirable), normality and homoscedasticity (Hair et al., 2010). The assumptions of factor analysis are vital and must be met for factor analysis to be regarded as appropriate. De Vet et al. (2005) state that the quality of results obtained from factor analysis depends largely on whether researchers conform to the assumptions underlying the principles and the methods that have been designed for appropriate application.

### 4.7.2.3 Performing the factor analysis

A critical decision in FA is the method of extraction. There are several factor analysis extraction methods to choose from. In the current study the researcher used Common factor analysis or Principal axis factoring (PAF). This method determines the least number of factors which can account for the common variance in a set of variables. It provides valuable insight into the structure of an assessment measure by explaining the covariance and correlation structure amongst the measured items (de Vet et al., 2005). In addition, principal axis factoring seeks the least number of factors which can account for the common variance shared by a set of variables (Garson, 2010; Hair et al., 2010). One would therefore refer to the communalities in order to assess the common variance. Communalities represent the proportion of variance in each variable which can be explained by all extracted factors together, and may be interpreted as the *reliability of the indicator*. In other words, a variable's communality is a measure of its shared or common variance among the variables as signified by the derived factors (Hair et al., 2010). Accordingly, Hair et al. (2010) advise that the communalities should be examined in order to ensure an adequate level is maintained after extraction.

It is important to note that PAF assumes that the sample is the population, so the results cannot be extrapolated beyond our particular sample, unless analysis using different samples reveals the same factor structure (Field, 2005). Since this is an initial validation of the KIDSCREEN-52 the researcher wishes to only apply the findings to the sample collected. Another important decision in FA revolves around the number of factors to retain for interpretation and further use (Hair et al., 2010). In deciding how many factors to extract, the researcher must blend a conceptual foundation (i.e. how many factors must be in the structure) with some empirical evidence (i.e. what amount of factors can be reasonably supported). In this study, after extraction, factors were retained for rotation based on a predetermined number of factors derived from the first research objective of the current study as well as prior research. This *a priori* criterion for retaining factors is reasonable when the researcher already knows how many factors to extract before undertaking the factor analysis (Hair et al., 2010). This approach, according to Hair et al. (2010) is justified if the researcher endeavours to obtain the same number of factors previously found.

Components were rotated obliquely using the direct oblimin rotation method. Factor rotations are utilised to discriminate between factors (Field, 2005) and facilitate and improve the interpretation of factors (Hair et al., 2010; Henson & Roberts, 2006). Hair et al. (2010) state that the primary outcome of rotating the factor matrix is to restructure the variance from preceding factors to later ones in order to achieve a more straightforward and theoretically meaningful factor pattern (Hair et al, 2010). With oblique rotation factors are allowed to correlate, which presupposes that the underlying factors are theoretically related (Field, 2005). Field (2005) suggests that if there are theoretical grounds for supposing that your factors might correlate, then factors should be rotated using direct oblimin rotation.

The next step relates to the choice of the matrix to utilise in the interpretation of the factor analysis. The appropriate matrix chosen for this study was the pattern matrix. When using oblique rotations, Costello and Osborne (2005) recommend that one should examine the pattern matrix for factor item loadings. The pattern matrix contains information about the unique contribution of a variable to a particular factor (Field, 2005). This matrix simplifies and clarifies the data structure and therefore renders interpretation easier. The researcher can thus assess the number of items loading on a factor, the values of the loadings as well as each items contribution to a factor(s) and the contribution of the factor(s) to the construct being measured. In addition, one can also identify if any variable has a cross-loading, i.e. a variable has more than one significant loading (Field, 2005) in which case the researcher can take any of a specified number of actions. Hair et al. (2010) suggests that the researcher may (1) either ignore problematic variables, interpret the solution as is and note the variables which are poorly presented in the factor solution; (2) evaluate variables for possible deletion; (3) use an alternative rotation method; (4) decrease/increase the number of factors retained; or modify the type of factor model used.

When a satisfactory factor solution has been derived, the final step is to assign some meaning to the factors (Hair et al., 2010). This involves careful interpretation of the pattern of factor loadings for a variable.

### 4.7.3    Assessing the internal consistency reliability of the KIDSCREEN-52

**Research Aim 2: To assess the internal consistency reliability of each of the sub-scales of the KIDSCREEN-52.**

The method used for examining the internal consistency reliability of the KIDSCREEN-52 was the statistical technique of Cronbach's alpha. This involves deriving an index, ranging from 0-1, based on the correlations of the indicators that comprise the measure (O'Leary-Kelly & Vokurka, 1998). Cronbach's alphas were calculated for the scale as a whole as well as on the original scales of the KIDSCREEN-52.

### 4.7.3.1 Reliability

There is a close relationship between validity and reliability. Kaplan and Saccuzzo (2005) maintain attempting to define validity of a measure will be pointless if the test is not reliable. The reliability of a measure refers to its capacity to produce consistent results. It is defined as a matter of whether a specific technique, applied repeatedly to the same object, would produce the same result each time (Babbie & Mouton, 2001). The concept reliability is based on two key considerations: (1) do items purportedly belonging to a scale actually assess a single construct, and (2) do scales measuring a single construct produce consistent estimates of that construct across multiple measurements (Switzer, Wisniewski, Belle, Dew & Schultz, 1999; Hair et al., 2010). Reliability is therefore essentially an assessment of the degree of consistency between multiple measurements of a variable (Hair et al., 2010). However, reliability not only pertains to the consistency or stability of a measure, but is also inversely related to the degree to which a measure is contaminated by random error. The score obtained by a test-taker on a measure comprises two components: an underlying "true" score and random error (Hammond, 2006). Thus, reliability is measured in terms of the proportion of true score variance in relation to observed score variance. Random error is always present to a certain degree and can constitute a major problem which can jeopardise the validity of research findings (O'Leary-Kelly & Vokurka, 1998).

The classical way of assessing reliability is the test-retest method, by which consistency is measured between the responses for an individual at two points in time (Blaikie, 2003; Hair et al., 2010). However, this method has a number of drawbacks which restricts its effectiveness as an estimate of reliability. Firstly, the testing circumstances may be different for the test-taker (e.g. anxiety and illness) and the physical environment (e.g. noise and venue) may also differ from that of the first application (Foxcroft & Roodt, 2005). Secondly, in the measurement of attitude, current information or new experiences, may lead to a shift in attitude (Singleton et al., 1993). Failing one or two subjects in a final exam for example may change one's attitude towards school. In addition, transfer effects, such as practice and memory, might impact on the second testing occasion (Foxcroft & Roodt, 2005).
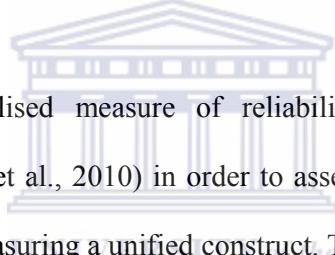
Another, more commonly utilised measure of reliability is to examine the internal consistency of test items (Hair et al., 2010) in order to assess the extent to which items are related to each other and are measuring a unified construct. The ensuing coefficient (normally Cronbach's alpha) is rooted in the homogeneity of the items, with a high alpha emerging when the items correlate well together (Hammond, 2006). Hammond (2006) maintains that in this way, reliability may be regarded as a manner of construct validation. Moreover, high internal consistency is regarded as a precondition for high validity (Kline, 1993).

The current study assessed the internal consistency reliability on the original scales of the KIDSCREEN-52 with Cronbach's alpha which, according to O'Leary-Kelly and Vokurka (1998), is the most popular method for assessing reliability. As mentioned previously, it entails deriving an index, ranging from 0-1, based on the correlations of the indicators that comprise the measure (O'Leary-Kelly & Vokurka, 1998).

The formula for Cronbach's alpha can be presented as follows:

$$\alpha = \frac{N^2\,\overline{\text{Cov}}}{\sum S^2_{\text{item}} + \sum \text{Cov}_{\text{item}}}$$

(Field, 2005)

For each item on our scale we can calculate two things, i.e. the variance within the item, and the covariance between a particular item and any other item on the scale (Field, 2005). It is generally agreed that the lower limit for Cronbach alpha is .70, although it may decrease to .60 in exploratory research (Hair et al., 2010). However, Blaikie (2003) states that the value of alpha is affected by the number of items in a scale; and alpha increases as the number of items in a scale increases. Overall, the researcher anticipates that the current study will yield relatively high reliability indices (above .70).
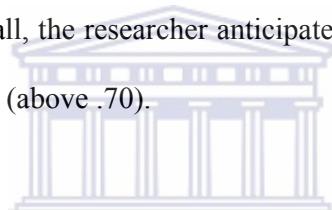
## 4.8 Ethical considerations

In order to protect the rights and welfare of participants (Terre Blanche, Durrheim & Painter, 2006) dealing appropriately with ethical issues is viewed as a vital component in any research which involves human subjects. Accordingly, participation in the main study was voluntary and voluntary informed consent and assent was obtained from parents and learners respectively, in addition to permission from teachers and the school. The current study was executed in conformity with the ethical code of professional conduct in social research. Permission and ethical clearance for the current study was granted by the University of the Western Cape Ethical Committee, and permission was also granted by the key investigator of the main study (Appendix D). As the current study forms part of the broader study, it made use of secondary data. Information was primarily anonymous with the only identifying data being age, gender, religion, area of residence and school. The data was handled with the utmost care; stored in a secure place and only handled by the researcher of the current study.

**4.9 Conclusion**

This chapter outlined the methodology of the current study and the procedures that were followed to achieve the aims and objectives of the study, namely to evaluate the reliability and construct validity of the KIDSCREEN-52 in a South African context. This chapter also described the types of analyses used in order to meet the aims of the study. The subsequent chapter illustrates and reports the results of these analyses.

UNIVERSITY *of the*
WESTERN CAPE

## CHAPTER 5

## RESULTS

### 5.1 Introduction

This chapter will illustrate the findings by means of a tabular representation of the results and thereafter interpret the analyzed data. It will, in particular, look at the quality of completion of the data, the analysis of compliance with specific assumptions in conducting EFA, the results of the EFA, and the Cronbach's alpha coefficients of the ten subscales of the KIDSCREEN-52. These interpretations of the data will pave the way for the discussion of the outcomes of these results in the subsequent chapter.

### 5.2 Quality of completion

A full range of score distributions was obtained across dimension scales of the KIDSCREEN-52. De Vet et al. (2005) maintain that in conducting factor analysis, it is vital to provide information regarding the number of missing values on each item as well as how the researcher dealt with the missing values. They state that this is necessary because an accumulation of missing values as a result of paired missings may seriously restrict the number of subjects which serves as a basis on which the variance covariance matrix is founded. They considered missing values of 25% in any item as the maximum. The data in table 6 below indicates how many participants answered all items in each KIDSCREEN-52 dimension as well as the missing data.

**Table 6:**

**Missing data**

| | N | Mean | Std. Deviation | Missing | |
|---|---|---|---|---|---|
| | | | | Count | Percent |
| Physical Activities and Health Index | 545 | 17.32 | 4.03 | 20 | 3.5 |
| Feelings Index | 553 | 22.68 | 4.95 | 12 | 2.1 |
| General Mood Index | 536 | 29.33 | 3.04 | 29 | 5.1 |
| About Yourself Index | 555 | 20.68 | 2.74 | 10 | 1.8 |
| Free Time Index | 548 | 17.84 | 5.12 | 17 | 3.0 |
| Family and Home Life Index | 547 | 22.42 | 5.94 | 18 | 3.2 |
| Money Matters Index | 552 | 9.38 | 3.52 | 13 | 2.3 |
| Friends Index | 547 | 22.45 | 5.28 | 18 | 3.2 |
| School and Learning Index | 548 | 21.27 | 5.10 | 17 | 3.0 |
| Bullying Index | 553 | 13.59 | 1.29 | 12 | 2.1 |

The overall proportion of scale-level missing data was 2.93%. The dimension that presented with the most missing data was "General Mood" (5.1%) and the dimension with the least missing data was "About Yourself" (1.8%). In this study, the missing values were acceptable since the proportion of scale level missing data was less than the maximum value of 25% (de Vet et al., 2005).

Missing data may be dealt with by: (1) replacing it with the mean (which may perhaps lead to significant results that may not otherwise be significant); (2) excluding cases listwise (i.e. any participant with missing data for any variable is excluded) or (3) excluding cases pairwise (i.e. a participant's data are only excluded from computations for which a datum is missing) (Field, 2005). Since excluding cases pairwise could result in estimates being all over the place, Field (2005) suggests that it is safest to exclude cases listwise if it does not result in a

massive loss of data. Accordingly, since these values were minimal and considered to be missing at random, cases were excluded from the analyses listwise.

## 5.3 Exploratory factor analysis

The 52 items of the KIDSCREEN measure were subjected to an exploratory factor analysis utilising principal axis factoring with oblique rotations using the direct oblimin rotation method. The factor analysis yielded a ten-factor solution based on *a priori* criterion for retaining factors.

### 5.3.1 Reporting the steps in the factor analysis

Reporting and being explicit about the steps in the analysis is essential to assist other researchers to interpret, corroborate and replicate results (de Vet et al., 2005). As a first step the factoriability of the data was assessed by means of the KMO measure of sampling adequacy and the Bartlett's test of sphericity. Thereafter, factors were rotated using direct oblimin rotation (step 2) because of the correlation among variables. To determine the number of factors to retain (step 3) an *a priori* criterion of 10 factors was specified. Using the pattern matrix for interpretation, the cut off score to determine factor loadings was .30 for retention of items, based on the given sample size (as specified by Hair et al., 2010). After a satisfactory factor solution was derived, the final step entailed assigning meaning to the factors by careful interpretation of the pattern of factor loadings for a variable.

### 5.3.2 Analysis of compliance with specific assumptions

According to Hair et al. (2010), before the factor analysis is performed the researcher must establish that a strong conceptual foundation provides support for the assumption that a structure exists. Firstly, factor analysis requires a normal distribution of the data.
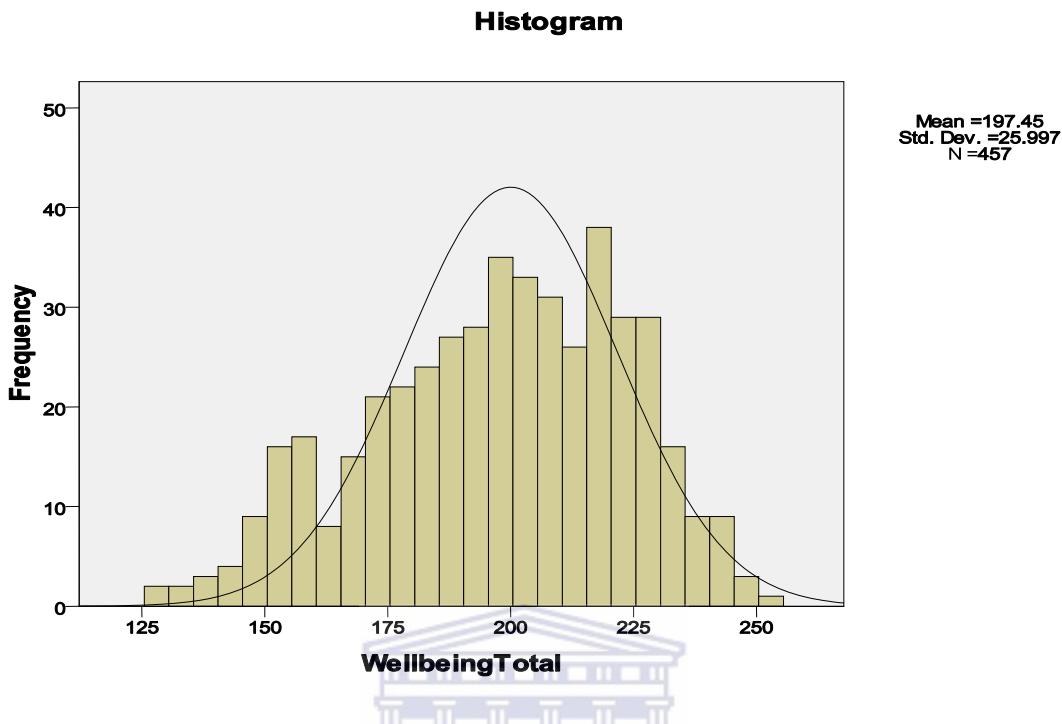
**Histogram**



**Figure 1: A histogram indicating the distribution of the total well-being scores**

It is clear from figure 1 above that the data displays a bell-shaped curve, indicating that our data are normally distributed.

As mentioned previously, the stability of factor analytic results rests heavily on sample size and larger samples produce more stable results. The sample size (n = 565) of the current study is therefore sufficient for factor analysis with more than 10 participants per variable.

Hair et al. (2010) highlight three approaches to screen the data matrix in order to ascertain whether it is suitable to apply factor analysis to the data. In order to justify the application of factor analysis, the first step is to assess the factoriability of the correlation matrix via a visual inspection of the correlations. A visual inspection of the correlation matrix (Appendix E) reveals that all the variables correlate with one another. A substantial number of correlations greater than .30 indicate that positive correlations do exist between the variables entered for

the factor analysis with the majority of correlations being significant (p < .05). In addition, the review also revealed that none of the items correlate very highly (.90 or greater) which diminishes the concern for singularity in the data. Since none of the correlation coefficients (Appendix E) exceed or come close to .90 we can safely say that multicollinearity is not a problem for this data set. It is thus clear that all items of the KIDSCREEN-52 correlate fairly well with all others and none of the correlation coefficients are particularly large, there is thus no need to consider eliminating any questions at this point of the analysis (Field, 2005).

The next approach to quantify the extent of intercorrelations among the variables and to assess the factoriability of the data is to ascertain whether the item intercorrelations comply with the criterion of sample adequacy required for EFA. The researcher conducted the Kaiser Meyer Olkin (KMO) test which is linked to the Measure of Sampling Adequacy (table 7).
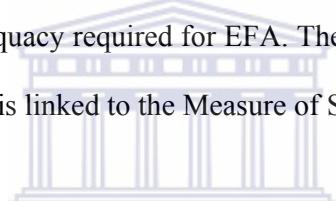
**Table 7:**

**KMO and Bartlett's Test**

| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | | .894 |
|---|---|---|
| **Bartlett's Test of of Sphericity** | Approx. Chi-Square | 9994.321 |
| | Df | 1326 |
| | Significance | .000 |

The KMO measure of sampling adequacy is a test of the amount of variance within the data that could be explained by factors (Brace, Kemp & Snelgar, 2003). Table 7 above indicates that there was sufficient intercorrelations between the variables for factor analysis with the KMO Measure of Sampling Adequacy at .894, thus >.50 and closer to 1, which falls within the range of being meritorious (Field, 2005). As a result, we can be confident that factor

analysis is appropriate for these data. In addition, since KMO is more than .50, these data are linearly distributed.

The Bartlett's test of Sphericity (table 7) provides the statistical significance that the correlation matrix has significant correlations among at least some of the variables (Hair et al., 2010). This tests the null hypothesis that the correlation matrix is an identity matrix (a matrix in which all the diagonal elements are 1 and all off diagonal elements are 0) (Field, 2005; Hair et al., 2010). The Bartlett's test of Sphericity (table 7) is highly significant with $p < .001$, which means that the $R$-matrix differs significantly from an identity matrix (Field, 2005). Thus, there are statistically significant correlations between some of the variables, which render the correlation matrix adequate for conducting a factor analysis.
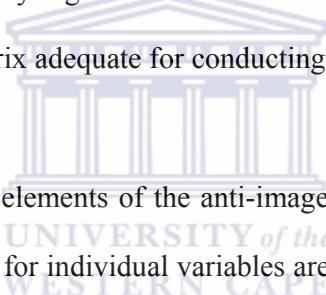
An examination of the diagonal elements of the anti-image correlation matrix (Appendix F) indicates that the KMO statistics for individual variables are well above the bare minimum of .50 for all the variables.  All the correlations are above .70, which is regarded as high (Field, 2005) with the exception of one, "About Yourself" item 3, with a value of .694.  Thus, all the variables can be included in the analysis. In addition, the off-diagonal elements represent the partial correlations between the variables and for a good factor analysis we want these correlations to be very small. All the off-diagonal values (Appendix F) of this data are relatively small.

Together, the statistically significant Bartlett's test of Sphericity, the MSA values for both overall test and each individual variable (anti-image matrices) provide evidence that there is sufficient intercorrelation and common variance between the variables, and that the current data is adequate to perform a factor analysis.

### 5.3.3 Deriving factors and assessing overall fit

The factors of the KIDSCREEN-52 were rotated by means of oblique rotations using the direct oblimin rotation method. As mentioned previously, when using oblique rotations, it is advisable that one should examine the pattern matrix for factor item loadings (Costello & Osborne, 2005). Using the pattern matrix for interpretation, the cut off score to determine factor loadings was .30 for retention of items, based on the sample size (n-565) (as specified by Hair et al., 2010).

Since the researcher already knew the number of factors to extract before undertaking the analysis, an *a priory* criterion was applied in the current study. This approach is justified as the current study attempted to extract the same number of factors that was previously found in European studies (Hair et al., 2010). The results obtained and illustrated in table 8 below, indicate the loadings of the different items on each factor.

**Table 8:**

**Factor Pattern Matrix**

|  | Factor | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
|  | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
| Family HomeLife4 | **.776** | | | | | | | | | |
| Family HomeLife6 | **.764** | | | | | | | | | |
| Family HomeLife2 | **.725** | | | | | | | | | |
| Family HomeLife3 | **.636** | | | | | | | | | |
| Family HomeLife1 | **.564** | | | | | | | | | |
| Family HomeLife5 | **.489** | | | | | | | | .216 | |
| School Learning1 | | **.748** | | | | | | | | |
| School Learning5 | | **.690** | | | | | | | | |
| School Learning2 | | **.684** | | | | | | | | |
| School Learning3 | | **.676** | | | | | | | | |

| Item | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| School Learning6 | **.633** | | | | | | | |
| School Learning4 | **.471** | | | | | | | |
| General Mood3 | | **.688** | | | | | | |
| General Mood4 | | **.531** | | | | | | |
| General Mood2 | | **.512** | | | | | | |
| General Mood5 | | **.470** | | | | | | |
| General Mood6 | | **.423** | | | | | | .269 |
| General Mood1 | .203 | **.398** | | | | | | |
| General Mood7 | | **.334** | | | | | | |
| Bullying2 | | .223 | | | | .201 | | |
| Friends4 | | | **.709** | | | | | |
| Friends5 | | | **.635** | | | | | |
| Friends6 | | | **.633** | | | | | |
| Friends3 | | | **.570** | | | .243 | | |
| Friends1 | | | **.465** | | | .262 | | |
| Friends2 | | | **.357** | | | | | |
| Bullying3 | | | .185 | | | | | |
| PhysicalActHealth4 | | | | **.787** | | | | |
| PhysicalActHealth3 | | | | **.708** | | | | |
| PhysicalActHealth2 | | | | **.645** | | | | |
| PhysicalActHealth5 | | | | **.467** | | | -.287 | |
| PhysicalActHealth1 | | | | **.350** | | | | |
| Money Matters1 | | | | | **.888** | | | |
| Money Matters3 | | | | | **.861** | | | |
| Money Matters2 | | | | | **.822** | | | |
| FreeTime3 | | | | | | **.586** | | |
| FreeTime2 | | | | | | **.579** | | |
| FreeTime4 | | | | | | **.555** | | |
| FreeTime5 | | | | | | **.496** | | |
| FreeTime1 | | | | | | **.457** | -.256 | |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Feelings2 | | | | | | | | **-.606** | | |
| Feelings3 | .205 | | | | | | | **-.582** | | |
| Feelings1 | | | | | | | | **-.469** | -.192 | |
| Feelings5 | | | | | | | | **-.682** | | |
| Feelings4 | | | | | | | | **-.597** | | |
| About Yourself1 | .276 | | | | | | | **-.355** | | |
| Feelings6 | | | | **.311** | | | | **-.315** | | |
| About Yourself3 | | | | | | | | | | **.624** |
| About Yourself4 | | | | | | | | | | **.596** |
| About Yourself5 | | | | | | | | | | **.481** |
| Bullying1 | | | | | | | | | | .224 |
| About Yourself2 | | | | | | | | | | .211 |

The *a priori* exploratory factor analysis extracted a ten factor structure which was derived from the hypothesised dimensions stipulated in the current study, and findings from earlier European studies (Ravens-Sieberer et al., 2005). The ten factors, which were stipulated for extraction, explained 56.91 % of the variance in the items. All the items loaded separately as expected, except the items relating to "Bullying", "About Yourself" and "Feelings" which deviated from these earlier findings.

The first factor to emerge was the "Parent relations and Home Life" factor, which comprised "Family and Home Life" items 1-6. All the items comprising the hypothesised dimension

loaded on the empirically derived factor. Loadings were high (ranging from .489 to .776) indicating a solid factor.

The second factor generated was the "School Environment" factor which consisted of "School and Learning" items 1-6, with all six hypothesised items also loading significantly on the factor. Loadings ranged from .471 to .748.

Similar findings were observed with factor 3 "Moods and Emotions" (which comprise General Mood items 1-7), factor 4 "Social Support and Peers" (which comprise "Friends" items 1-6), factor 5 "Physical Well-being" (which consist of "Physical Activities and Health items 1-5), factor 6 "Financial Resources" (which comprise "Money Matters" items 1-3) and factor 7 "Autonomy" (which consist of "Free Time" items 1-5).These factors also had all the hypothesised items loading on the factors.  These factors all had 3 or more items loading with values higher than .50 which indicates solid factors.

Whilst the last three factors also had significant loadings of .30 and above, items comprising the hypothesised dimensions deviated from the expected item loadings. Two scales, "Feelings" and "About Yourself" items divided into three scales. The eighth factor "Psychological Well-being" was well defined with 3 of the hypothesised items ("Feelings" items 1-3) loading saliently on the factor (.606, .582, .469) indicating a stable factor. However, two of the hypothesised items ("Feelings" items 4 and 5) relating to this dimension loaded on factor 9 "Self Perception" and one item ("Feelings" item 6) cross-loaded significantly onto two different factors, i.e. factor 4 "Social Support and Peers" and factor 9 "Self Perception". Factor 9 contained only one of the hypothesised items ("About Yourself" item 1) which loaded significantly on this factor.

Four "About Yourself" items loaded on factor 10 "Social Acceptance (Bullying)" factor with all but one loading significantly. Only one of the hypothesised "Bullying" items loaded on this factor. The other two "Bullying" items also loaded insignificantly on factor 3 and 4 respectively.

The substantial deviations from the hypothesised loadings and the one significant cross-loading of the abovementioned three questions with other domains, can be explained in terms of the wording of questions or the way the questions were posed, the close relationship between the various dimensions tapped by the questions and the language factor as is discussed in the following chapter.
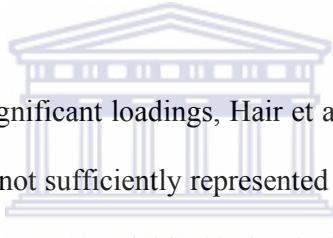
After having identified all the significant loadings, Hair et al. (2010) advise the researcher to search for any variables that are not sufficiently represented by the factor solution. They state that this is achieved by identifying any variable (s) that lack at least one significant loading, or by examining each variable's communality. It is clear from table 7 above that "Bullying" items are poorly presented in the factor solution. None of the three bullying items had significant loadings; only one loaded insignificantly on the factor 10; one loaded on moods and emotions and one loaded on social support and peers. The communalities (table 8) for the bullying items were also low (< .50) indicating that the "Bullying" dimension is not sufficiently represented by the factor solution.

Communalities represent the proportion of variance in each variable which can be explained by all extracted factors together and may be interpreted as the *reliability of the indicator* (Field, 2005; Hair et al., 2010). In other words, a variable's communality is the estimate of its shared or common variance among the variables as represented by the derived factors (Field,

2005). The communalities are viewed to assess whether the variables meet acceptable levels of explanation. Whilst it is conventional to interpret the communalities before looking at the item loadings on the different factors, Garson (2010) explains that communalities must be interpreted in relation to the interpretability of the factors. Hence, the researcher of the current study chose to first look at the interpretability of the factors and then at the communalities. Table 9 below displays the communalities of the 52 variables both before and after extraction.

**Table 9:**

**Communalities**

|  | Initial | Extracted |
|---|---|---|
| PAH1 | .243 | **.198** |
| PAH2 | .450 | .467 |
| PAH3 | .479 | .500 |
| PAH4 | .536 | .622 |
| PAH5 | .438 | .422 |
| Feelings1 | .489 | .485 |
| Feelings2 | .532 | .563 |
| Feelings3 | .535 | .566 |
| Feelings4 | .559 | .587 |
| Feelings5 | .615 | .702 |
| Feelings6 | .535 | .470 |
| GenMood1 | .310 | **.304** |
| GenMood2 | .309 | **.299** |
| GenMood3 | .345 | .440 |
| GenMood4 | .339 | **.334** |
| GenMood5 | .310 | **.313** |
| GenMood6 | .315 | **.331** |
| GenMood7 | .248 | **.207** |

| | | |
|---|---|---|
| AYourself1 | .479 | .461 |
| Ayourself2 | .328 | **.284** |
| Ayourself3 | .294 | **.365** |
| Ayourself4 | .280 | **.365** |
| Ayourself5 | .232 | **.243** |
| FreeTime1 | .534 | .494 |
| FreeTime2 | .589 | .606 |
| FreeTime3 | .516 | .527 |
| FreeTime4 | .547 | .533 |
| FreeTime5 | .506 | .497 |
| FhomeLife1 | .536 | .517 |
| FhomeLife2 | .642 | .665 |
| FhomeLife3 | .617 | .613 |
| FhomeLife4 | .667 | .715 |
| FhomeLife5 | .467 | .480 |
| FhomeLife6 | .622 | .656 |
| Mmatters1 | .669 | .741 |
| Mmatters2 | .647 | .701 |
| Mmatters3 | .673 | .765 |
| Friends1 | .536 | .494 |
| Friends2 | .352 | **.285** |
| Friends3 | .614 | .551 |
| Friends4 | .566 | .605 |
| Friends5 | .528 | .435 |
| Friends6 | .499 | .465 |
| School1 | .579 | .599 |
| School2 | .532 | .553 |
| School3 | .521 | .484 |
| School4 | .406 | **.364** |
| School5 | .538 | .577 |
| School6 | .477 | .429 |
| Bullying1 | .286 | **.234** |

| Bullying2 | .255 | **.247** |
|-----------|------|----------|
| Bullying3 | .269 | **.200** |

The values in the "extraction" column in table 9 indicate the proportion of each variable's variance that can be explained by the retained factors. Variables with high values are well represented in the common factor space, whereas variables with low values are poorly represented. In table 9 above we can see that the value for "Physical Activities and Health" item 1 (19.8%) is particularly low. In relation to the proportion of variance explained by the underlying factors, we can say that after we extracted factor 1, the amount of variance left unexplained is 100 – 19.8 = 80.2%.

In general, communalities indicate for which measured variables the factor analysis is working best and least well. Hair et al. (2010) maintain that variables with communalities less than .50 do not have sufficient explanation. They further state that variables with low values are not well represented in the common factor space, whilst variables with high values are well represented. Thus, we can say that the factor analysis works best for "Money Matters" item 3 (.765) but not as well for "Physical Wellbeing" item1 (.198). However, if we look at the pattern matrix (table 8) we can see that this item has a significant loading (.350) on factor 5. Garson (2010) gives a good explanation for this discrepancy. He states that a communality of .75 appears high, but is actually meaningless if the factor on which the variable is loaded is not interpretable, though it usually is. Similarly, a communality of .25 appears low but may be meaningful if the item contributes to a well-defined factor as can be seen from the communality and the factor loading of "Physical Activity and Health" item 1. Thus, what is important is not the communality coefficient per se, but rather the degree to which the item plays a part in the interpretation of the factor, although this role is often greater when communality is high.

We can see from table 8 (Pattern matrix) that all our items contribute meaningfully to the factors, with the exception of the three "bullying" items which also have low communality values (< .50). Thus, these items do not have sufficient explanation. Although the hypothesised "Bullying" items did not load as expected, table 10 below indicates that the "Bullying" scale displayed high internal consistency (.81), indicating that in an 11 factor solution these items may load consistently on an additional factor.

**Table 10:**

**Reliability statistics of the KIDSCREEN-52scales**

| Scale | Alpha |
|---|---|
| Physical Activities and Health Index | .79 |
| Feelings Index | .76 |
| General Mood Index | .80 |
| About Yourself Index | .79 |
| Free Time Index | .76 |
| Family and Home Life Index | .77 |
| Money Matters Index | .79 |
| Friends Index | .78 |
| School and Learning Index | .80 |
| Bullying Index | .81 |

In order to assess the overall fit of the model we need to look at the differences between the observed correlations and the correlations based on the model which is quoted in the lower half of the reproduced matrix labelled *Residual* (Appendix E) (Field, 2005). For a good model, these values should be small, less than .05 and the percentage of non-redundant residuals with absolute values >.05 should be less than 50%. The percentage of non-

redundant residuals with absolute values >.05 is 5% as there are 70 non-redundant residuals. This indicates that our model is a good model and since most residuals are small, it is not necessary to extract further factors.

### 5.3.4 Labelling the factors

When an appropriate factor solution has been obtained, the next step is to assign some meaning to the factors. This step entails identifying the variables with the greatest contribution to a factor and assigning a name to represent the factor's conceptual meaning (Hair et al., 2010). Thus, the researcher will place greater emphasis on those variables with higher loadings when assigning a name to a factor. The label or name is intuitively produced by the researcher based on its suitability to represent the underlying dimensions of a specific factor (Hair et al., 2010). Thus, naming the factors is based on the subjective opinion of the researcher, but in this study it will also be guided by the theory.
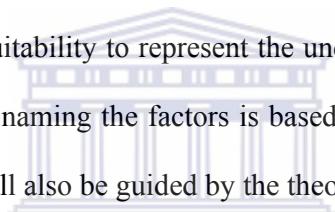
As mentioned in the theoretical discussion in section 2.5, if a measure is designed to assess a certain number of dimensions defined by means of a specified succession of item sets, then the dimensions defined in this manner should emerge in the investigation of a representative empirical problem (Derogatis & Melisaratos, 1983). The expected dimensions to emerge from this analysis was outlined in section 2.7 and these are: physical well-being, psychological well-being, moods and emotions, self perception, autonomy, parent relations and home life, social support and peers, school environment, social acceptance (bullying) and financial resources. In addition, the current analysis of the internal structure of the KIDSCREEN-52 reveals the extent to which the relationships between test items and test components are consistent with the construct (HRQoL) on which the postulated test score interpretations are established.

Since we need at least three variables to load significantly on a factor in order for it to be stable, we will name the factors based on the three most significant loadings only. Thus, each factor can be named as follows based on the variable with significant loadings:

**Table 11:**

**Factor names, items and loadings**

| Factor Number | Factor Name | Item number and item | Variable Loading |
|---|---|---|---|
| Factor 1 | **Parent Relations and Home Life** | 4. Have your parents had enough time for you? <br> 6. Have you been able to talk to your parent(s) when you wanted to? <br> 2. Have you felt loved by your parent(s)? | .776 <br> .764 <br><br> .725 |
| Factor 2 | **School Environment** | 1. Have you been happy at school? <br> 5. Have you enjoyed going to school? <br> 2. Have you got on well at school? | .748 <br> .690 <br> .684 |
| Factor 3 | **Moods and Emotions** | 3. Have you felt so bad that you didn't want to do anything? <br> 4. Have you felt that everything in your life goes wrong? <br> 2. Have you felt sad? | .688 <br> .531 <br> .512 |
| Factor 4 | **Social Support and Peers** | 4. Have you and your friends helped each other? <br> 5. Have you been able to talk about everything with your friends? <br> 6. Have you been able to rely on your friends? | .709 <br> .635 <br><br> .633 |
| Factor 5 | **Physical Well-being** | 4. Have you been able to run well? <br> 3. Have you been physically active? <br> 2. Have you felt fit and well? | .787 <br> .708 <br> .645 |
| Factor 6 | **Financial Resources** | 1. Have you had enough money to do the same things as your friends? <br> 3. Do you have enough money to do things with your friends? <br> 2. Have you had enough money for your expenses? | .888 <br><br> .861 <br><br> .822 |
| Factor 7 | **Autonomy** | 3. Have you had enough opportunity to be outside? <br> 2. Have you been able to do things that you want to do in your free time? <br> 4. Have you had enough time to meet friends | .586 <br> .579 <br><br> .555 |
| Factor 8 | **Psychological Well-being** | 2. Have you felt pleased that you are alive? <br> 3. Have you felt satisfied with your life? <br> 1. Has your life been enjoyable? | -.606 <br> -.582 <br> -.469 |

| Factor 9 | **Emotional Self perception** | 5. Have you felt cheerful? (Feelings5) | -.682 |
| | | 4. Have you been in a good mood? (Feelings4) | -.597 |
| | | 1. Have you been happy with the way you are (About Yourself1) | -.355 |
| Factor 10 | **Social Self Perception** | 3. Have you been worried about the way you look? | .624 |
| | | 4. Have you felt jealous about the way other girls and boys look? | .596 |
| | | 5. Would you like to change something about your body? | .481 |

The first seven factors in table 11 above were based on theory as hypothesised in chapter 2. The last three factors were named based on the researcher's subjective deliberation. Factors 8 and 9 have negative loadings. According to DeCoster (1998) the interpretation of negative loadings can be facilitated by multiplying variable loadings by -1 to get a positive value and renaming the factor with its polar opposite. However, in this instance, although factor 8 had three loadings, the researcher decided to keep the hypothesised label of the original scale, i.e. "Psychological Well-being" as psychological well-being occurs on a continuum and can therefore range from negative to positive. There is thus no need to use DeCoster's line of reasoning in this instance.

Since a negative self-image may affect an individual's mood and feelings of happiness, the items that loaded on factor 9 appear to be measuring perceptions of affect or emotions. Factor nine was therefore named "Emotional self-perception". The items on factor ten all refer to self-perception about physical appearance so, it was named "Social Self Perception"

## 5.4 Reliability

As a measure of internal consistency of the KIDSCREEN-52, Cronbach's alphas were calculated for the scale as a whole as well as for each of the ten subscales. Field (2005) highlights that tests which have normally distributed scores, are likely to have higher

reliability estimates than tests with positively or negatively skewed distributions. Thus, one must take into account the distribution of scores in relation to the estimate of the $\alpha$ as part of the interpretation. The descriptive statistics and estimates of internal consistency for the total scale and the various dimensions of the KIDSCREEN-52 are reported in table 12 and 13 below.

**Table 12:**

**Reliability statistic KIDSCREEN-52 total**

| Cronbach's Alpha | N of Items |
|---|---|
| .80 | 10 |

**Table 13:**

**Descriptive statistics and estimates of reliability of the KIDSCREEN-52scales**

| Scale | Mean | SD | Variance | Alpha |
|---|---|---|---|---|
| Physical Activities and Health Index | 180.04 | 4.02 | 580.15 | .79 |
| Feelings Index | 174.61 | 4.95 | 502.27 | .76 |
| General Mood Index | 168.17 | 3.06 | 613.64 | .80 |
| About Yourself Index | 176.95 | 2.74 | 599.77 | .79 |
| Free Time Index | 179.55 | 5.10 | 501.30 | .76 |
| Family and Home Life Index | 174.97 | 6.02 | 473.36 | .77 |
| Money Matters Index | 187.89 | 3.52 | 586.64 | .79 |
| Friends Index | 174.67 | 5.15 | 526.56 | .78 |
| School and Learning Index | 176.36 | 5.13 | 549.13 | .80 |
| Bullying Index | 183.87 | 1.30 | 660.50 | .81 |

The analysis revealed satisfactory to good Cronbach's alpha coefficients for the whole scale (table 12) as well as for all the individual subscales (table 13). The reliability index is .80 (table 12), which can be regarded as very reliable. As can be seen in table 13 above, all the alpha coefficients were above .70 for all the subscale scores, ranging from .76 to .81, indicating satisfactory to good internal consistency. These high levels of $\alpha$ engender confidence in the reliability of the KIDSCREEN-52 within a South African context.

## 5.5 Conclusion

The current chapter provided a tabular presentation of the statistical analysis conducted to examine the factorial validity and internal consistency reliability of the KIDSCREEN-52 in a South African context. The following chapter will discuss these findings, highlight the limitations present in conducting the study as well as offer some recommendations to improve the validity of the KIDSCREEN-52 for South African populations.
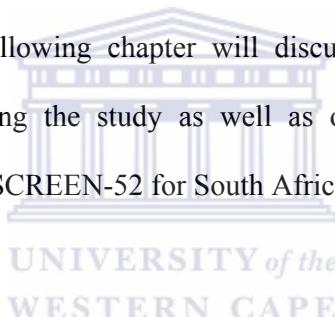
# CHAPTER 6

# DISCUSSION AND CONCLUSION

## 6.1 Introduction

The current study sought to establish the first psychometric properties of the KIDSCREEN-52 in order to validate this instrument for South African children and adolescents. The overall aim of the present study was to examine the validity and reliability of the KIDSCREEN-52 in a South African context. In order to achieve these objectives, it examined the factor structure of the KIDSCREEN-52 and explored the internal consistency reliability of the 10 subscales of this measure. The overall findings, based on the results yielded from the analysis, will be discussed in this chapter. The discussion will also focus around the implications of using the KIDSCREEN-52 without examining its psychometric properties further and make some suggestions to improve its validity in a South African context.
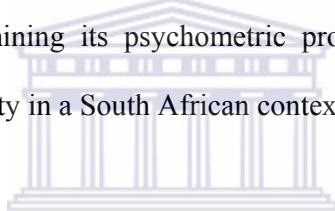
At the same time, the discussion encourages a critical reflection on the importance of issues surrounding cross-cultural testing, test adaptation and the role of language in the assessment process. This chapter will provide a brief summary of the core arguments present in the paper, the limitations in the current study and provide suitable recommendations for future research. The chapter will terminate with final concluding remarks.

## 6.2 Discussion of results

The results of the present study provide initial evidence on the psychometric properties of the KIDSCREEN-52 within a South African context. The overall findings were scrutinized to support the aims of the current study. Various psychometric aspects were examined, including feasibility, factoriability of the data, factorial structure and reliability.

Feasibility was assessed by examining the proportion of missing items within the KIDSCREEN-52 scales. Overall, all the scales had fewer than 5% missing responses with the exception of the "General Mood Index" which had 5.1% missing responses. Since these missing values were less than the maximum value (25%) stated by de Vet et al. (2005), it was deemed acceptable. Since these values were minimal and considered to be missing at random, they were excluded from the analysis.

### 6.2.1 Internal structure and construct validity

Exploratory Factor Analysis using principal axis factoring with oblique rotation extracted ten factors as predetermined by means of *a priori* criterion for retaining factors. The amount of variance explained by the ten factors was 56.91%. Overall this model had a good fit and all the assumptions of factor analysis were met. However, the empirically derived factor structure deviates slightly from the structure previously identified by the European KIDSCREEN group. The first 7 factors all had three or more salient loadings with values higher than .50 indicating solid factors. Also, all the items comprising these seven hypothesised dimensions loaded on the empirically derived factors as expected.

The eighth factor (Psychological well-being) had 3 of the hypothesised items loading significantly (> .40), which renders the factor valid. Even though these loadings were negative, the researcher decided to keep the hypothesised label, i.e. "Psychological Well-being" as psychological well-being occurs on a continuum and can therefore range from negative to positive.

In addition, two of its hypothesised items (Feeling items 4 and 5) "Have you been in a good mood?" and "Have you felt cheerful?" loaded on factor 9 (Self Perception) along with

"About Yourself" item 1, "Have you been happy with the way you are?" Since a negative self-image may affect an individual's mood and feelings of happiness, these items appear to be conceptually similar which may be why these feelings items split from the eighth factor. Hence, the items that loaded on factor 9 appear to be measuring perceptions of affect or emotions and the factor was therefore named "Emotional self-perception".

However, they could have been misconstrued to refer to fun and excitement instead of the intended self-perception. Also, the last feelings item (item 6; "Have you had fun?") had significant cross-loadings on factor 4 (Social Support and Peers) as well as on factor 9 (Self Perception) which indicates that this item could be measuring a totally different construct.
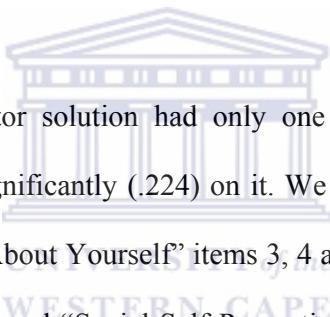
The final factor in the ten-factor solution had only one (Bullying item 1) of the three hypothesised items loading insignificantly (.224) on it. We can therefore conclude that they are not useful here. However, "About Yourself" items 3, 4 and 5 loaded significantly (>.040) on factor 10. This factor was named "Social Self Perception" as the items on this factor all refer to self-perception about physical appearance. Although this factor can still be retained based on the significant loadings of items 3, 4 and 5, it is not clear why these items loaded on this factor. Further investigation into these items might provide a clearer picture as to why this has occurred.

The substantial loadings of items from these three questions with other domains can also be explained by the formulation of question-items used in the KIDSCREEN-52. For example, the question regarding "self-perception", "Have you been worried about the way you look?" can be linked to the "Social acceptance" (Bullying) factor. Documented research (Baron, Byrne & Branscombe, 2006) has shown that how others view you, impacts on how you view

yourself. So this item could very well have been interpreted as "Have you been worried about how others think you look?" and thus, loading on the domain of social acceptance. It seems like "self-perception" has two factors: "emotional self-concept" and "physical self-perception".

Moreover, participants might have had a problem in understanding the questions due to the way questions were formulated, or language could have played a role. The fact that the KIDSCREEN-52 was standardised and normed on a European English first language speaking sample, and that a significant percentage of the participants of the current study's first language was not English (25.5% had Xhosa as a first language and 13.8% had Afrikaans as a first language), is a cause for concern. As discussed in section 3.4, language difficulties can be an obstacle in assessment (Foxcroft & Roodt, 2005). The level of language comprehension may place a lower limit on the appropriateness of a particular questionnaire to certain test-takers (Matza et al., 2004). Moreover, an individual's grade level is not an accurate reflection or indication of his/her reading ability or literacy level (Wasserman, Maja & Wright, 2010).Thus, one cannot presume that a learner in grade 9 or 10 is able to fully understand a test administered in English if their first language is Afrikaans or Xhosa. This is especially true in a South African context where English literacy levels and reading comprehension abilities are low among certain communities.

Accordingly, since the KIDSCREEN-52 was administered in English, without any adaptations in terms of language and cultural concerns, this could have played a significant role in the results obtained from the current study. Instrument as well as sample characteristics could have some bearing on the probable occurrence of bias (Van de Vijver, 1998). Thus, it is suggested that future research assess issues of bias and differential item

functioning (DIF) across the two countries and run an item difficulty analysis to see if any of the items were challenging to the participants. DIF analysis is a method used to identify items that function differently among different groups, and therefore help monitor the validity and fairness of questionnaires (Lin & Rogers, 2005).

Although there are certain deviations in the empirical structure from the hypothesised factor structure, there is more similarity than divergence between the two since the first 7 factors were reproduced with no deviation from the hypothesised structure and the other three factors were also significant. Results from this analysis therefore add weight to the construct validity of the KIDSCREEN-52 in a South African context.

### 6.2.2 Internal consistency reliability

Cronbach alpha coefficients were computed to estimate the internal consistency reliability for each of the factor based scales of the KIDSCREEN-52 as well as the overall summary score. Alpha coefficients of .70 or higher were regarded as satisfactory. Overall, the KIDSCREEN-52 demonstrated a strong internal consistency within a South African context in terms of both the total scale (.80) and individual domain scores (> .70). In the ten-domain structure, $\alpha$ scores were satisfactory to good (ranging from .76 to .81) in all the domains, as shown in table 11. This indicates that the KIDSCREEN-52 is an internally reliable tool for the assessment of HRQoL in South African populations. These findings are similar to the α coefficients reported for the KIDSCREEN-52 by the European KIDSCREEN group (.77 to .89) in all participating European countries except for in the case of one scale "Bullying" (Social acceptance) which was found to be below .70 in one county (France) (Ravens-Sieberer et al., 2005). The findings of the current study are also analogous to those of the Korean study conducted by Hong et al. (2006) with the translated K-KIDSCREEN-52. They

also found alpha coefficients of .76 and above for all dimensions. This indicates that in a South African context, the KIDSCREEN-52 performs as well as it did in the European and Korean context, with comparable reliability coefficients.

## 6.3 Limitations of results

1. Since interpretation of results in EFA is largely dependent on reflective researcher judgment, it is dependent on the researcher's knowledge and expertise which may be a limitation in the current study.

2. Utilising secondary data restricts the researcher in the analysis by the original objectives of the research and the researcher is not able to control for data collection errors.

3. Even though factor analysis can illuminate the underlying meaning of scores obtained from the use of a particular measure, Goodwin (2000) states that an overdependence on factor analysis in validation research can lead to a restricted body of findings. Since this study is only the first step in the validation of the KIDSCREEN-52 in a South African context, this study merely serves as a basis for further investigation into the construct validity of this measure in this context.

4. A further limitation in the current study is that the researcher did not investigate problematic items further in terms of item difficulty and issues of fairness and bias which would have shed more light on the way these items loaded on the factors. However, this is beyond the scope of the present study.

**6.4 Recommendations**

1. Based on the results of the current study, items should be examined for bias. An analysis of differential item functioning (DIF) should be performed to identify whether problem items identified in the KIDSCREEN-52 are understood in a South African cultural context, whether language played a role in the responses to item (i.e. the way the items were answered), whether the wording of items was clear and whether the type of questions asked was familiar to the respondents.

2. Since the "Bullying" scale displayed high internal consistency, indicating that in an 11 factor solution these items may load consistently on an additional factor, it is recommended that further research explore an 11 factor solution.

3. It is also recommended that future research should focus on an item analysis in order to determine whether the items serve the purpose for which it was designed, to determine how difficult an item is and what the shortcomings of the item are.

4. It is also recommended that after examining bias and DIF, that confirmatory factor analysis be performed to determine whether the conclusions on the factor structure of the KIDSCREEN-52 in a South African context would replicate the European structure.

5. Further research needs to also include the study of the instrument's external relations to other variables, from a convergent and a discriminant framework.

**6.5 Concluding remarks**

The importance of psychosocial aspects of HRQoL in children and adolescence echoes

throughout the literature and highlights the need for research in this area with this population. As mentioned by Ravens Sieberer et al., (2005) until now, items and dimensions are relevant to children and adolescents of participating European countries and it still needs to be seen whether this also holds true for children and adolescents in other countries. The current study took the initial step to comply with this recommendation by assessing the construct validity of the KIDSCREEN-52 within a South African context. The results provide strong evidence that social self-perception and emotional self-perception are two different factors and not one dimension.

This initial psychometric analysis provides validity evidence as to the ability of the KIDSCREEN-52 for sound measurement with sufficient psychometric properties with some further suggested analysis and possible minor adaptation. This study provided preliminary evidence to support the internal consistency reliability and validity of the KIDSCREEN-52 within a South African context. However, as recommended, further testing in other samples is needed to assess the cross-cultural validity of the KIDSCREEN-52 as well as to generalise to the population at large.

This study highlights the serious implications of using tests in different cultural contexts without taking into account the first/home language of participants. Although, there is a growing awareness of bias and equivalence in South Africa, many researchers do not evaluate their instruments in order to make it culturally and linguistically appropriate. Foxcroft (2002) highlights that few bias studies have been undertaken in South Africa and very little effort has been invested in adapting mono-cultural, Westernized tests for the African context. She further emphasises that it is essential to raise awareness among assessment practitioners about

the unacceptability of using tests without having undertaken bias studies, adapting and re-norming them, and without having established 'local' psychometric properties.

As far as Messick's notion of the consequential basis of construct validity, researchers should be cognisant of the fact that certain tests can have grave implications on individuals' lives. QoL instruments are increasingly used in the health sector to make important decisions about patient's health needs and it is therefore vital that such instruments are reliable and valid for use in such populations. A test that endeavours to measure HRQoL, but in reality measures something else, can lead to distorted interpretations of results.

UNIVERSITY *of the*
WESTERN CAPE

# REFERENCES

Abbott, R.A., Ploubidis, G.B., Huppert, F.A., Kuh, D., Wadsworth, M.E.J. & Croudace, T.J. (2006). Psychometric evaluation and predictive validity of Ryff's psychological well-being items in a UK birth cohort sample of women. *Health and Quality of Life Outcomes.* [On-line]. Accessed 10 January 2010. Available: http://www.hqlo.com/content/4/1/76.

American Educational Research Association (AERA), American Psychological Association (APA) & National Council on Measurement in Education (NCME). (1999). *Standards for educational and psychological testing.* Washington: American Educational Research Association.
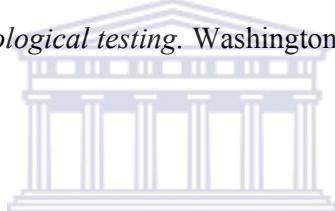
Anastasi, A. & Urbina, S. (1997). *Psychological Testing.* (7th ed.). USA: Prentice Hall International.

Arnold, R., Ranchor, A.V., Sanderman, R., Kempen, G.I.J.M., Ormel, J. & Suurmeijer, T.P.B.M. (2004). The relative contribution of domains of quality of life to overall quality of life for different chronic diseases. *Quality of Life Research, 13,* 883-896.

Babbie, E. (2004). *The practice of social research.* (10th ed.). Belmont: Wadsworth/Thomson Learning.

Babbie, E. & Mouton, J. (2001). *The practice of social research.* (9th impression)**.** South Africa: Oxford University Press.

Baron, R.A., Byrne, D. & Branscombe, N.R. (2006). *Social psychology.* (9th ed.).New York: Pearson.

Bisegger, C., Cloetta, B., von Rueden, U., Abel, T. & Ravens-Sieberer, U. (2005). Health related quality of life: gender differences in childhood and adolescence. *Soz Präventimed, 50,* 281-291.

Blaikie, N. (2003). *Analyzing quantitative data.* London: Sage Publications Ltd.

Bookwalter, J.T. & Dalenberg, D. (2004). Subjective Well-being and Household factors in South Africa. *Social Indicators Research, 65,* 333-353.

Bullinger, M., Brűt, A.L., Erhart, M. & Ravens-Sieberer,U. (2008). Psychometric properties of the KINDL-R questionnaire: results of the BELLA study. *European Child & Adolescent Psychiatry, 17*(1), 125-132.

Bullinger, M., Schmidt, S., Petersen, C. & Ravens-Sieberer, U. (2006). Quality of life – evaluation criteria for children with chronic conditions in medical care. *Journal of Public Health, 14,* 343-355.

Burkhardt, C.S., Anderson, K.L., Archenholtz, B. & Hägg, O. (2003). The Flanagan Quality of Life Scale: Evidence of Construct Validity. *Health and Quality of Life Outcomes, 1*(59). [On-line]. Accessed 27 May 2010. Available: http://www.hqlo.com/content/1/1/59.

Cavallo, F., Zambon, F., Borraccino, A., Ravens-Sieberer, U., Torsheim, T. & Lemma, P. (2006). Girls growing through adolescence have a higher risk of poor health. *Quality of Life Research, 15*(10), 1577-1585.

Claes, C., Van Hove, G., van Loon, J., Vandevelde, S. & Schalock, R.L. (2009). Quality of Life Measurement in the Field of Intellectual Disabilities: Eight Principles for Assessing Quality of Life –Related Personal Outcomes. *Social Indicators Research, 98*(1), 61-72.

Costello, A.B. & Osborne, J.W. (2005). Best Practices in Exploratory Factor Analysis: Four recommendations for getting the most from your Analysis. *Practical Assessment Research and Evaluation, 10*(7). [On-line]. Accessed 27 May 2010. Available: http://pareonline.net/getvn.asp?v=108cn=7.

Cramer, J.A., Westbrook, L.E., Devinsky, O. Perrine, K., Glassman, M.B. & Camfield, C. (1999). Development of the Quality of Life in Epilepsy Inventory for Adolescents: The QOLIE-AD-48. *Epilepsia, 40*(8), 1114-1121.

Cronbach, L.J. & Meehl, P.E. (1955). Construct Validity in Psychological Tests. *Psychological Bulletin, 52,* 281-302.

Currie, C., Gabhainn, S.N. & Godeau, E. (2009). The Health Behaviiiour in School-aged

    Children: WHO Collaborative Cross-National (HBSC) Study: Origins, concept, history

    and development 1982-2008. *International Journal Public Health, 54,* S131-S139.


Dale, A., Arber, S. & Procter, M. (1988). *Doing secondary analysis.* London: Unwin Hyman.


Davis, E., Nicolas, C., Waters, E., Cook, K., Gibbs, L., Gosch, A. & Ravens-Sieberer, U.

    (2007). Parent-proxy and child self-reported health-related quality of life: using

    qualitative methods to explain the discordance. *Quality Life Research, 16,* 863-871.


DeCoster, J. (1998). *Overview of factor analysis.* [On-line]. Accessed 15 March 2010.

    Available: http://www.stat-help.com/notes.html.
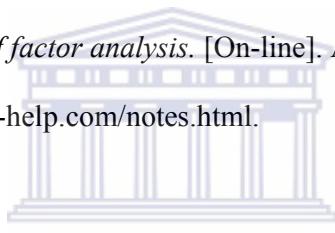

De Klerk, G. (2008). Cross-cultural testing. In Born, Foxcroft & Butter (Eds.), *Online*

    *Readings in Testing and Assessment,* International Test Commission. [On-line].

    Accessed 15 July 2010. Available: http://www.intest.com.org/Publications/ORTA.php.


Della, L.J., De Joy, D.M., Goetzel, R.M., Ozminkowski, R.J. & Wilson, M.G. (2008).

    Assessing Management Support for Worksite Health promotion: Psychometric

    Analysis of the Leading by Example (LBE) Instrument. *Am J Health Promotion, 22*(5),

    359-367.


Derogatis, L.R. & Melisaratos, N. (1983). The Brief Symptom Inventory: an introductory

    report. *Psychological Medicine, 13,* 595-605.

Detmar, S.B., Bruil, J., Ravens-Sieberer, U., Gosch, A., Bisegger, C. (2006). The use of focus groups in the development of the KIDSCREEN HRQOL questionnaire. *Quality of Life Research, 15,* 1345-1353.

De Vet, H.C.W., Adèr, H.J., Terwee, C.B. & Pouwer, F. (2005). Are factor analytical techniques used appropriately in the validation of health status questionnaires? A systematic review on the quality of factor analysis of the SF-36. *Quality of Life Research,14,* 1203-1218.

Duarte, P.S., Ciconelli, R.M. & Sesso, R. (2005). Cultural adaptation and validation of the "Kidney Disease and Quality of Life – Short Form (KDQOL-SF™ 1.3)" in Brazil. *Braz J Med Biol Res, 38*(2)*,* 261-270.

Eiser, C. & Morse, R. (2001). Can parents rate their child's health-related quality of life? Results of a systematic review. *Quality of Life Research, 10,* 347-357.

Erhart, M., Ellert, U., Kurth, B. & Ravens-Sieberer, U. (2009). Measuring adolescents' HRQoL via self-reports and parent-proxy reports: an evaluation of the psychometric properties of both versions of the KINDL-R instrument. *Health and Quality of Life Outcomes,7*(77). [On-line]. Accessed 15 January 2010. Available: http://www.hqlo.com/content/7/1/77.

Erhart, M., Ottova, V., Gaspar, T., Jericek, H., Schnohr, C., Alikasifoglu, M., Morgan, A., Ravens-Sieberer, U. and the HBSC Positive Health Focus Group. (2009). Measuring

mental health and well-being of school-children in 15 European countries using the KIDSCREEN-10 index. *Int J Public Health, 54,*160-166.

Erhart, M., Ravens-Sieberer, U., O'Dickinson, H. & Colver, A. (2009). Rasch Measurement Properties of the KIDSCREEN Quality of Life Instrument in Children with Cerebral Palsy and Differential Item Functioning between children with and without Cerebral Palsy. *International Society for Pharmacoeconomics and Outcomes Research (ISPOR), 12*(5), 782-792.

Fallowfield, L (2009). *What is quality of life?* [On-line]. Accessed 03 December 2009. Available: www.whatisitseries.co.uk.
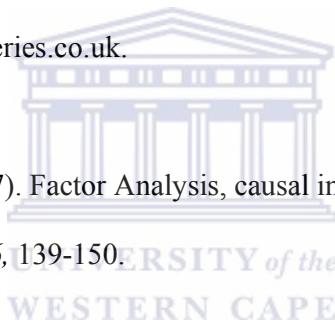
Fayers, P.M. & Hand,.D.J. (1997). Factor Analysis, causal indicators and quality of life. *Quality of Life Research, 6,* 139-150.

Field, A. (2005). *Discovering statistics using SPSS.* (2nd ed.). London: Sage Publications.

Fife-Shaw, C. (2006). Introduction to structural equation modelling. In Breakwell, Hammond &Fife-Schaw (Eds.), *Research methods in psychology* (pp.444-464). London: Sage Publications.

Foxcroft, C.D. (2002). Ethical Issues Related to Psychological testing in Africa: What I have learned (so far). *International Association for Cross-cultural Psychology, 4*(5). [On-line]. Accessed 03 June 2010. Available: http://www.ac.wwu.edu/culture/foxcroft.htm.

Foxcroft, C. & Roodt, G. (2005). *An introduction to psychologicalassessment in the South African context.* (2<sup>nd</sup> ed.). South Africa: Oxford University Press.

Frisch, M.B., Clark, M.P., Rouse, S.V., Rudd, M.D., Paweleck, J.P., Greenstone, A. & Kopplin, D.A. (2005). Predictive and Treatment Validity of Life Satisfaction and Quality of Life Inventory. *Assessment, 12*(1), 66-78.

Fryback, D.G. (2003). Workshop to Develop a Research Agenda and Research Resources for Health Status assessments and Summary Health Measures. March 26-28. [On-line]. Accessed 20 April 2010. Available: http://www.cdc.gov/nchs/about/IAWG/NCHS_Report.pdf.

Garson, D. (2010). *Factor Analysis: Stat notes: Topics in multivariate analysis.* [On-line]. Accessed 03 June 2010. Available: http://faculty.chass.ncsu.edu/garson/PA765/factor.htm.

Gentile, S., Jouve, E., Dussol, B., Moal, V., Berland, Y. & Sambuc, R. (2008). Development and validation of a French patient-based health-related quality of life instrument in kidney transplant : the ReTransQol. *Health and quality of Life Outcomes, 6*(78). [On-line]. Accessed 10 January, 2010. Available: http://www.hqlo.com/content/6/1/78.

Ghiselli, P.E., Campbell, E.E. & Zedeck, S. (1981). *Measurement theory for the behavioural sciences.* USA: W.H. Freeman and Company.

Goldbeck, L., Schmitz, T.G., Besier, T., Herschbach, P. & Henrich, G. (2007). Life satisfaction decreases during adolescence. *Quality Life Research, 16,* 969-979.

Goodwin, L.D. (2000). The Role of Factor Analysis in the Estimation of Construct Validity. *Measurement in physical education and exercise science, 3*(2), 85-100.

Government Gazette. (1998). *Republic of South Africa.* (Vol. 400, no. 19370). Cape Town,19 October 1998. [On-line]. Accessed 30 January 2010. Available : www.info.gov.za/View/DownloadFileAction?id=70714.

Greef, M., Uys, L.R., Wantland, D., Makoae, L., Chirwa, M., Dlamini, P., Kohi, T.W., Mullan, J., Naidoo, J.R., Cuca, Y. & Holzemer, W.L. (2010). Perceived HIV stigma and life satisfaction among persons living with HIV/Aids infection in five African countries. *International Journal of Nursing Studies, 47*(4), 475-486.

Guillemin, F., Bombardier, C. & Beaton, D. (1994). Cross-cultural adaptation of health-Related Quality of life measures: literature review and proposed guidelines. *Psychological Abstracts,* 8263569.

Hair, J.F., Black, W., Babin, B.J. & Anderson, R.E. (2010). *Multivariate data analysis.* (7th ed.). N.J.: Prentice Hall.

Hammond, S. (2006). Using Psychometric Tests. In Breakwell, Hammond & Fife-Schaw (Eds.), *Research Methods in Psychology* (pp.182-209). London: Sage Publications.

Hays. R D. (2003). Thoughts on Assorted Issues in Health Related Quality of Life

Assessment. *To Develop a research agenda and research resources for Health Status Assessment and Summary Health Measures: Workshop Report,* 111-117. [On-line]. Accessed 20 April 2010. Available:

http://www.cdc.gov/nchs/about/IAWG/NCHS_Report.pdf.


Henson, R.K. & Roberts, J.K. (2006). Use of Exploratory Factor Analysis in Published

Research: Common Errors and Some Comment on Improved Practice. *Educational and Psychological Measurement, 66*(3), 393-416.


Herdman, M., Rajmil, L., Ravens-Sieberer, U., Bullinger, M., Power, M., Alonso, J. And the

European Kidscreen and Disabkids groups. (2002). Expert consensus in the development of a European health-related quality of life measure for children and adolescents: a Delphi study. *Acta Paeditrica, 91,* 1385-1390.


Herjanic, B., Herjanic, M., Brown, F. & Wheatt, T. (1975). Are children reliable reporters?

*Journal of Abnormal Child Psychology,3*(1), 41-48.


Hong, J.H., Jeon, S.S., Lee, H.M., Choi, Y., Kim, S. & Choi, H.Y. (2006). The Functional

Assessment of Cancer Therapy-Prostate (FACT-P) Scales in Men with Prostate Cancer: reliability and Validity of the Korean Version. *J Korean Med Science, 21,* 295-299.


Hubley, A.M. & Palepu, A. (2007). Injection Drug User Quality of Life Scale (IDUQOL):

Findings from a content validation study. *Health and Quality of Life Outcomes, 5,* 46.

Janssens, L., Gorter, J.W., Ketelaar, M., Kramer, W.L.M. & Holtslag, H.R. (2008). Health related quality of life measures for long-term follow-up in children after major trauma. *Quality Life Research, 17,* 701-713.

Jelsma, J. & Ferguson, G. (2004). The determinants of self-reported health-related quality of life in a culturally and socially diverse South African Community. *Bulletin of the World Health Organisation, 82*(3), 206-212.

Jelsma, J., MacLean, E., Hughes, J., Tinise, X. & Darder, M. (2005). An investigation into the health related quality of life of individual's living with HIV who are receiving HAART. *AIDS Care, 17*(5). 579-588.

Jirojanakul, P., Skevington, S.M. & Hudson, J. (2003). Predicting young children's quality of life. *Social Science and Medicine, 57,* 1277-1288.

Jooste, M.J.L. (2001). *Introduction to psychological testing.* Johannesburg: Rand Afrikaanse Universiteit.

Judd, C.M., Smith, E.R. & Kidder, L.H. (1991). *Research methods in social relations.* USA:Harcourt Brace Javanovich College Publishers.

Kaplan, R.M., Bush, J.W. & Berry, C.C. (1976). Health Status: Types of validity and the index of Well-being" *Health Services Research.* 478-507. [On-line]. Accessed 10 January 2010. Available: www.ncbi.nlm.nih.gov/pmc/articles/pmc10719/pdf/hsresearch00553-0150.pdf.

Kaplan, R.M. & Saccusso, D.P. (2005). *Psychological testing : Principles, applications, and issues.* USA : Thomson.

Keenaghan, C. & Kilroe, J. (2008). A study on the quality of life tool KIDSCREEN for children and adolescents in Ireland : Results of the KIDSCREEN National Survey 2005. *The National Children's Strategy Research Series.*[On-line]. Accessed 18 June 2010. Available : omc@health.gov.ie.

Korevaar, J.C., Merkus, M.P., Jansen, M.A.M., Dekker, F.W., Boeschoten, E.W. & Krediet, R.T. (2002). Validation of the KDQOL-SF      : A dialysis-targeted health measure. *Quality of Life Research, 11,* 437-447.
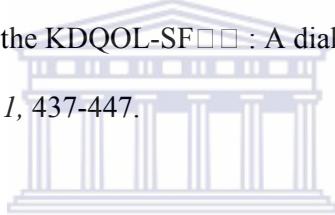
Lin, J. & Rogers, W.T. (2005). *Validity of the Simultaneous Approach to the Development to of Equivalent Achievement Tests in English and French (Satge II).*Alberta : Centre for Research in Applied Measurement and Evaluation, The University of Alberta.

Lix, L.M., Metge, C. & Leslie, W.D. (2009). Measurement equivalence of osteoporosis-specific and general quality-of-life instruments in Aboriginal and non-Aboriginal women. *Quality Life Research, 18,* 619-627.

Lundqvist, C., Rugland, E., Clench-Aas, J., Bartonova, A. & Hofoss, D. (2010). Children are reliable reporters of common symptoms: results from a self-reported symptom diary for primary school children. *Acta Paediatrica, 99*(7), 1054-1059.

Makiwane, M. & Kwizera, S. (2009). Youth and Well-Being: A South African Case Study. *Social Indicators Research, 91,* 223-242.

Markus, K.A. (1998). Validity, Facts, and Values Sans Closure: Reply to Messick, Reckase, Moss and Zimmerman. *Social Indicators Research, 45,* 73-82.

Marshall, M.N. (1996). Sampling for qualitative research. *Family Practice, 13*(6). 522-525.

Matza, L.S., Swensen, A.R., Flood, E.M., Secnik, K. & Leidy, N.K. (2004). Assessment of Health-Related Quality of Life in Children: A Review of Conceptual, Methodological, and Regulatory Issues. *Value in Health, 7*(1), 79-92.

Messick, S. (1989). Validity. In Linn (Ed.), *Educational measurement* (pp.13-104). New York: Macmillan Publishing Company.

Messick, S. (1998). Test Validity: A matter of consequence. *Social Indicators Research, 45*, 35-44.

Michel, G., Bisegger, C., Fuhr, D.C. & Abel, T. (2009). Age and gender differences in health-related quality of life of children and adolescents in Europe: a multilevel analysis. *Quality of Life Research, 18,* 1147-1157.

Møller, V. (2001). Monitoring Quality of Life in cities: Durban case. *Development Southern Africa, 18*(2), 217-238.

Møller, V. (2005). Resilient or resigned? Criminal Victimisation and Quality of Life in South Africa. *Social Indicators Research, 72,* 263-317.

Møller, V. (2007). Quality of Life in South Africa: The first ten years of democracy. *Social Indicators Research, 81,* 181-201.

Mosam, A., Vawda, N.B., Gordhan, A.H., Nkwanyanat, N. & Aboobaker, J. (2005). Quality of life issues for South Africans with acne vulgaris. *Clinical and Experimental Dermatology, 30,* 6-9.

Mouton, J. (2001). *How to succeed in your Master's and Doctoral studies: A South African guide and resource book.* Pretoria: Van Schaik.
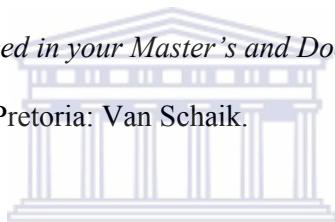
Muldoon, M.F., Barger, S.D., Flory, J.D. & Manuck, S.B. (1998). Education and debate: What are quality of life measurements measuring? *BMJ, 316*, 542-545.

Myant, K.A. & Williams, J.M. (2005). Children's concepts of health and illness: Understanding of contagious illnesses, non-contagious illnesses and Injuries. *Journal of Health Psychology, 10*(6), 805-819.

O'Connor, R. (1993). Issues in the Measurement of Health-Related Quality of Life. *NHMRC National Centre for Health Program Evaluation*. Australia:  University of New South Wales. [On-line]. Accessed 08 December 2009. Available: http://www.rodoconorassoc.com.

O'Connor, R. (2004). *Measuring quality of life in health.* London: Churchill Livingstone.

O'Hare, T. (1997). Measuring Problem Drinking in First Time Offenders: Development of the College Alcohol Problem Scale (CAPS). *Journal of Substance Abuse Treatment, 14*(4). 383-387.

Okun, M.A. & Stock, W.A. (1987). The Construct Validity of Subjective Well-Being Measures: An Assessment via Quantitative Research Syntheses. *Journal of Community Psychology, 15,* 481- 491.

O'Leary, B. (2007). Changes in the Quality of Life of Durban's people. *Social Indicators Research, 81,* 357-373.

O'Leary-Kelly, S.W. & Vokurka, R.J. (1998). The empirical assessment of construct validity. *Journal of Operations Management, 16,* 387-405.

Onwuegbuzie, A.J., Daniel, L.G. &Collins, K.M.T. (2009). A meta-validation model for assessing the score-validity of student teaching evaluations. *Qual Quant, 43,* 197-209.

Orton, R.E. (1998). Sameul Messick's Consequential Validity. *Philosophy of Education.* [On-line]. Accessed 20 May 2010. Available: http://www.ed.uiuc.edu/eps/PES-Yearbook/1998/orton.

Panepinto, J.A., Hoffman, R.G. & Pajewski, N.M. (2009). A psychometric evaluation of the PedsQL™ Family Impact Module in parents of children with sickle cell disease. *Health*

*and Quality of Life Outcomes.* [On-line]. Accessed 15 January 2010. Available: http://www.hqlo.com/content/7/1/32.

Papalia, D.E., Olds, S.W. & Feldman, R.D. (2004). *A child's world: Infancy through adolescence.* New York: McGraw-Hill.

Pesudovs, K., Burr, J.M., Harley, C. & Elliot, D.B. (2007). The Development, Assessment, and Selection of Questionnaires. *Optometry and Vision Science, 84*(8). 663-674.

Pflug, J. (2009). Folk Theories of Happiness: A Cross-cultural Comparison of Conceptions of Happiness in Germany and South Africa. *Social Indicators Research, 92,* 551-563.

Phaladze, N.A., Human, S., Dlamini, S.B., Hulela, E.B., Hadebe, I.M. & Sukati, N.A. (2005). Quality of Life and the concept of "Living Well" with HIV/AIDS in Sub-Saharan Africa. *Journal of Nursing Scholarship, 37*(2), 120-126.

Pohlmann, J.T. (2004). Use and Interpretation of Factor Analysis in The Journal of Educational Research: 1992-2002. *The Journal of Educational Research, 18*(1), 14-22.

Rachiotis, G., Behrakis, P.K., Vasilou, M. & Yfantopoulos, J. (2006). Quality of life and smoking among industrial workers in Greece. *La Medicina del lavoro, 97*(1), 44-50.

Rai, Y., Dutta, T. & Gulati, A.K. (2010). Quality of life of HIV-Infected People Across Different Stages of Infection. *Journal of Happiness Studies, 11,* 61-69.

Rajmil, L., Herdman, M., De Sanmamed, M.F.,  Detmar, S., Bruil, J., Ravens-Sieberer, U., Bullinger, M., Simeoni, M., Auquier, P. and the KIDSCREEN Group. (2004). Generic Health-related Quality of Life Instruments in Children and Adolescents: A Qualitative Analysis of Content. *Journal of Adolescent Health, 34*(1), 37-45.

Ranwick, R. (n.d.). *Quality of Life*. Unpublished on-line notes. Quality of Life Research Unit, University of Toronto. [On-line]. Accessed 05 Decemeber 2009. Available: http://www.utoronto.ca/qol/

Ravens-Sieberer, U. Auquier, P., Erhart, M., Tebe, C., Bisegger, C., Dürr,, W., Cloetta, B. Czemy, L., Mazur, J., Czimbalos, A., Tountas, Y., Hagquist, C., Kilroe, J., & the KIDSCREEN Group. (2007). The Kidscreen-27 quality of life measure for children and adolescents: psychometric results from a cross-cultural survey in 13 European countries. *Quality of Life Research, 16*(8), 1347-1356.

Ravens-Sieberer, U. & Bullinger, M. (1998). Assessing health-related quality of life in chronically ill children with the German KINDL: first psychometric and content analytical results. *Quality of Life research, 7*, 399-407.

Ravens-Sieberer, U., Erhart, M., Gosch, A. & Wille, N. (2008). Mental Health of Children and Adolescents in 12 European Countries – Results from the European KIDSCREEN Study. *Clinical Psychology and Psychotherapy, 15*, 154-163.

Ravens-Sieberer, U., Erhart, M., Wille, N., Wetzel, R. Nickel, J. & Bullinger, M. (2006). Generic Health Related Quality of Life Assessment in Children and Adolescents. *Pharmacoeconomics, 24*(12), 1199-1220.

Ravens-Sieberer, U., Gosch, L., Rajmil, L., Erhart, M., Bruil, J., Duer, W., Auquier, P., Power, M., Abel, T., Czemy, L., Mazur, J., Czimbalmos, A., Tountas, Y., Hagquist, C., Kilroe, J. and the European KIDSCREEN Group. (2005). The KIDSCREEN-52 Quality of Life Measure for Child and Adolescent Development and first results from a European Survey. *Expert Rev Pharmacoecon Outcomes Res, 5*(3), 353-364.

Ravens-Sieberer, U., Gosch, L., Rajmil, L., Erhart, M., Bruil, J., Power, M., Duer, W., Auquier, P., Cloetta, B., Czemy, L., Mazur, J., Czimbalmos, A., Tountas, Y., Hagquist, C., Kilroe, J. and the KIDSCREEN Group. (2008). The KIDSCREEN-52 quality of life measure for children and adolescents: psychometric results from a cross-cultural survey in 13 European countries. *PubMed, 11*(4). 645-58.

Ravens-Sieberer, U., Schmidt, S., Gosch, A., Erhart, M., Petersen, C. & Bullinger, M. (2007). Measuring subjective health in children and adolescents: Results of the European KIDSCREEN/DISABKIDS Project. *Psychosoc Med, 4,* 1-17.

Ravens-Sieberer, U., Torsheim, T., Hetland, J., Volleberg,W., Cavallo, F., Jericek, H., Alikasifoglu, M., Välimaa, R., Ottova, V., Erhart, M and the HBSC Positive Health Focus Group. (2009). Subjective health symptom load and quality of life of children and adolescents in Europe. *International Journal Public Health, 54,* S151-S159.

Rebok, G., Riley, A., Forrest, C., Starfield, B., Green, B., Robertson, J. & Tambor, E. (2001). Elementary school-aged children's reports of their health: a cognitive interviewing study. *Quality of Life Research, 10*(1), 59-70.

Richards, R., O'Leary, B. & Mutsonziwa. (2007). Measuring Quality of Life in informal settlements in South Africa. *Social Indicators Research, 81,* 375-388.

Riley, A.W. (2004). Evidence that school age children can self-report on their health. *Ambul Pediatrica, 4*(4), 371-376.

Robitail, S., Simeoni, M., Erhart, M., Ravens-Sieberer, U., Bruil, J., Auquier, P. and the European Kidscreen Group. (2006). Validation of the European Proxy KIDSCREEN-52 Pilot Test Health-Related Quality of Life Questionnaire: First Results. *Journal of Adolescent Health,39,* 596.e1-596.e10.

Robitail, S., Ravens-Sieberer, U., Simeoni, M.C., Rajmil, L., Bruil, J., Power, M., Duer, W., Cloetta, B., Czemy, L., Mazur, J., Czimbalmos, A., Tountas, Y., Hagquist, C., Kilroe, J. & Auquier, P. – the KIDSCREEN Group. (2007). Testing the structural and cross-Cultural Validity of the KIDSCREEN-27 quality of life questionnaire. *Quality of Life Research,16.* 1335-1345.

Sapin, C., Simeoni, M., Khammar, M.E., Antoniotti, S. & Auquier, P. (2005). Reliability and validity of the VSP-A, a health-related quality of life instrument for ill and healthy adolescents. *Journal of Adolescent Health, 36,* 327-336.

Savahl, S., Isaacs, S., September, R. & Koch, E. (2009). *An investigation into the impact of exposure to community violence and hope on children's perceptions of well-being.* Cape Town: University of the Western Cape. In press.

Sechrest, L. (2005). Validity of Measures Is No Simple Matter. *HSR: Health Services Research, 40* (5), 1584-1604.

Singletary, L. (2005). Issues that affect youth quality of life in Lyon County: Results of a community situational analysis. [On-line]. Accessed 29 June 2010. Available: http://www.unce.unr.edu/publications/files/cd/2005/fs0505.pdf.
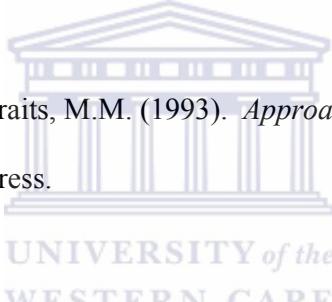
Singleton, R.A., Straits, B.C. & Straits, M.M. (1993). *Approaches to social research.* New York: Oxford University Press.

Skevington, S.M., Lofty, M. & O'Connell, K.A. (2004). The World Health Organization's WHOQOL-BREF quality of life assessment: Psychometric properties and results of the international field trial: A Report from the WHOQOL Group. *Quality of Life Research, 13,* 299-310.

Spieth, L.E. & Harris, C.V. (1996). Assessment of Health-Related Quality of Life in Children and Adolescents: An Integrative Review. *Journal of Pediatric Psychology, 21*(2), 175-193.

Springer, A.E., McQueen, A., Quintanilla, G., Arrivillaga, M. & Ross, M.W. (2009). Reliability and validity of the Student perception Perceptions of School Cohesion Scale

in a sample of Salvadoran secondary school students. *BMC international Health and Human Rights*.[On-line]. Accessed 10 January 2010. Available: http://www.biomedcentral.com/1472-698X/9/30.

Stewart, D.W. (1984). *Secondary research: information sources and methods.* London: Sage Publications Ltd.

Stewart-Brown, S., Tennant, A., Tennant, R., Platt, S., Parkinson, J. & Weich, S. (2009). Internal construct validity of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS): a Rasch analysis using data from the Scottish Health Education Population Survey. *Health and Quality of Life Outcomes, 7,* 15. [On-line]. Accessed 10 January 2010. Available: http://www.hqlo.com/content/7/1/15.

Strauss, M.E. & Smith, G.T. (2009). Construct Validity: Advances in Theory and Methodology. *Annual Rev Clin Psychol,* 5, 1-25.

Switzer, G.E., Wisniewski, S.R. Belle, S.H., Dew, M.A. & Schultz, R. (1999). Selecting developing, and evaluating research instruments. *Soc Psychology Psychiatr Epidemiol, 34,* 399-409.

Tebe, C., Berra, S., Herdman, M., Aymerich, M., Alonso, J. & Rajmil, L. (2008). Reliability and validity of the Spanish version of the KIDSCREEN-52 for child and adolescent population. *Medicina Clinica, 130*(17), 650-654.

Tennant, R., Hiller, L., Fishwick, R. & Platt, S. (2007). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS) development and UK validation. *Health and Quality of Life Outcomes, 5*(63). [On-line] Accessed 10 January 2010. Available: http://www.hqlo.com/content/5/1/63.

Terre Blanche, M., Durrheim, K. & Painter, D. (2006). *Research in practice: Applied methods for the social sciences.* (2nd ed.). Cape Town: University of Cape Town Press.

Torres, C.S., Paiva, S.M., Vale, M.P., Pordeus, I.A., Ramos-Jorge, M.L., Oliviera, A.C. & Allison, P.J. (2009). Psychometric properties of the Brazilian version of the Child Perceptions Questionnaire (CPQ) – short forms. *Health and Quality of Life Outcomes.* [On-line]. Accessed 10 January 2010. Available: http://www.hqlo.com/content/7/1/43.

Upton, P., Lawford, J. & Eiser, C. (2008). Parent-child agreement across child health-related quality of life instruments: a review of the literature. *Quality of Life Research, 17,* 895-913.

Van de Vijver, F.J.R. (1998). Towards a Theory of Bias and Equivalence. In Janet A. Harkness (ed.). *Cross-cultural Survey Equivalence.* Zuma-Nachrichten Spezial, nr.3. Manheim: ZUMA (Zentrum Für UmFragen, Methoden und Analysen). 41-62. [On-line]. Accessed 20 December 2009. Available: http://ww.gesis.org/fileadmin/upload/.../**zuma_nachrichten_spezial**/zn**spezial**3.pdf

Van de Vijver, A.J.R. & Rothman, S. (2004). Assessment in Multicultural Groups: The South African Case. *SA Journal of Industrial Psychology, 30*(4), 1-7.

Varni, J.W., Limbers, C.A. & Burwinkle, T.M. (2007). How young can children reliably and validly self-report their health-related quality of life?: An analysis of 8,591 children across age subgroups with the PedsQL™ 4.0 Generic Core Scales. *Health and Quality of Life Outcomes, 5*(1). [On-line]. Accessed 10 January 2010. Available: http://www.hqlo.com/content/5/1/1.

Von Reuden, U., Gosch, A., Rajmil, L., Bisegger, O. & Ravens-Sieberer, U. (2006). Socioeconomic determinants of health related quality of life in childhood and adolescence: results from a European study. *Journal of Epidemiology and Community Health, 60*(3), 130-135.

Walsh, W.B. & Betz, N.E. (2001). *Tests and assessment.* (4th ed.). New Jersey: Prentice Hall.

Wasserman, Z., Maja, T.M. & Wright, S.C.D. (2010). Assessment of English literacy level of patients in primary services in Tshwane, Gauteng Province: Part 2. *Health SA Gesondheid, 15*(1). [On-line]. Accessed 20 September 2010. Available: http://www.hsag.co.za/index.php/HSAG/article/viesfile/469/488.

Welman, C., Kruger, F. & Mitchell, B. (2005). *Research methodology.* South Africa: Oxford University Press.

Westaway, M.S. (2006). A longitudinal investigation of satisfaction with personal and environmental Quality of Life in an informal South African Housing settlement, Doornkop, Soweto. *Habitat International, 30*(1), 175-189.

Westaway, M.S., Olorunju, S.A.S. & Rai, L.C. (2007). Which personal quality of life domains affect the happiness of older South Africans. *Quality Life Research, 16*. 1425-1438.

Wilkins, A.J., O'Callaghan, M.J., Najman, J.M., Bor, W., Williams, G.M. & Shuttleworth, G. (2004). Early Childhood Factors Influencing Health-Related Quality of Life in Adolescents at 13 years. *Journal of Paediatrics and Child Health, 40*(3), 102-109.

World Health Organisation. (1948).Constitution of the World Health Organisation: Basic Document. Geneva, Switzerland, World Health Organisation.
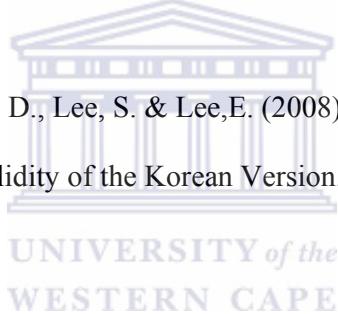
Yi, S.W., Kim, H., Lim, K., Bang, D., Lee, S. & Lee,E. (2008). The Behçet's Disease Quality of Life: Reliability and Validity of the Korean Version. *Yonsei Med Journal, 3*(5), 698-704).

Zullig, K.J., Valois, R. & Drane, J.W. (2005). Adolescent distinctions between quality of life and self-rated health in quality of life research. *Health and Quality of Life Outcomes, 3* (64). [On-line]. Accessed 10 January 2010. Available: http://www.hqlo.com/content/5/1/1.