

## Evaluating the construct validity of the KIDSCREEN-52 Quality of Life questionnaire in a South African context

**Naiema Taliep** (Medical Research Council /University of South Africa)

**Maria Florence** (University of the Western Cape)

**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 which was developed and standardised in Europe. The current study is part of a broader study conducted in the Western Cape, which used the KIDSCREEN-52 to explore the influence of exposure to community violence on the subjective HRQoL of a sample of South African adolescents. This study aimed to investigate the reliability and construct validity of the KIDSCREEN-52 in a South African context. The broader study employed stratified interval criterion sampling to select 565 Grade 9 learners, aged 14-18. Participants were selected from six public schools in areas specified by the South African government as comprising key nodal areas in terms of crime in the Western Cape. The dataset for the current study comprised all participants ( $N=565$ ) of the primary study. As the initial step in validation of the KIDSCREEN-52 in South Africa, the current study examined its factor structure 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 extracted 10 factors as identified by previous studies, with some deviation in the loadings of the last three factors. Items of two scales ("Feelings" and "About Yourself") divided into three scales, and "Bullying" items were not sufficiently presented in the factor solution. Internal consistency of the measure was shown to be acceptable to good, with Cronbach's alpha values ranging from 0.76 to 0.81 for the 10 scales.

**Keywords** Assessment; construct validity; exploratory factor analysis; KIDSCREEN-52; Quality of Life; reliability; South Africa.

This study was directed at examining the reliability and construct validity of the KIDSCREEN-52 Health Related Quality of Life (HRQoL) measure in a South African context. In South Africa, it is legislated (Employment Equity Act 55 of 1998) that the use of any psychometric instrument is permissible only when it has been shown to be scientifically valid for respondents from various cultures (Government Gazette, 1998). The disreputable roots of psychological testing seen in the inhumane treatment of mentally challenged people and the use of measures to show superiority of one race over another, served as an impetus to social and ethical considerations in the development and administration of psychological measures (Anastasi & Urbina,

1997; Foxcroft & Roodt, 2005). Without validation, tests may lead to biased interpretations for particular cultural groups, which can have grave consequences for people.

The development of instruments geared towards positive mental health and well-being, or HRQoL, is experiencing a renewed emphasis. However, a lack of appropriate measures renders this field under-researched (Hu, Stewart-Brown, Twigg & Weich, 2007). A dire need exists for measures that will permit HRQoL assessment (Robitail et al., 2006) with child and adolescent populations on a local and an international level. Only 13% of all HRQoL research publications are related to children, and only 9% to the testing of research instruments (Ravens Sieberer & Bullinger, 1998). The absence of a suitable assessment measure in a particular country compels researchers to develop new instruments, to adapt existing instruments, or to “export” an existing, almost always Euro-American, measure to the non-Western world, and establish its psychometric properties in the new context (Van de Vijver & Rothman, 2004).

Savahl, Isaacs, September and Koch (2009) explored the subjective HRQoL of a sample of South African adolescents from historically disadvantaged areas in the South Metropole of the Western Cape. They used the KIDSCREEN-52 in conjunction with the Children’s Hope Scale (Snyder et al., 1997) and the Recent Exposure to Violence scale (Singer, Anglin, Song & Lunghofer, 1995) to investigate the impact of hope as well as exposure to community violence on children’s perception of well-being (HECVW). The current study is an extension of the HECVW study as it examines the psychometric properties of the KIDSCREEN-52 in a South African context.

Though HRQoL instruments have gained prominence, a lack of consensus remains as to what constitutes Quality of Life (QoL). This ubiquitous concept has various philosophical, political and health-related dimensions (Fallowfield, 2009). It is often used interchangeably with health and well-being, as well as life-satisfaction (Kaplan, Bush & Berry, 1976; Goldbeck, Schmitz, Besier, Herschbach & Henrich, 2007). QoL is also a subjective notion, derived from the perceived impact that events and experiences have on an individual’s health and well-being. Gill and Feinstein (1994) emphasise that the absence of a unique definition for QoL highlights the importance of clearly defining the term whenever constructing or using a QoL instrument. QoL is generally defined 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 (WHOQOL, 1994). The World Health Organisation (WHO, 1948) defines health as a state of total physical, mental, and social well-being, and not simply the absence of disease. The construct HRQoL can therefore be described as a multi-dimensional psychological concept that encompasses functioning and well-being in the physical, psychological and social or emotional dimensions of life (Fallowfield, 2009; Ravens-Sieberer et al., 2006).

The KIDSCREEN-52 has been developed as a standardised instrument that can be applied in paediatric and healthy populations to assess the subjective HRQoL of children and adolescents (Ravens-Sieberer et al., 2005). It is a cross-culturally applicable measure developed along various cross-cultural approaches. It was

developed simultaneously in a number of European countries, and contains country-specific as well as multi-cultural aspects (Ravens-Sieberer et al., 2005). However, these cross-cultural validations occurred primarily in Europe, and the suitability for its cross-cultural use beyond Europe needs to be established.

The development of the KIDSCREEN-52 was based on literature reviews, focus group discussions and expert consultation (Delphi method) (Detmar, Bruil, Ravens-Sieberer, Gosch, Bisegger & The European KIDSCREEN Group, 2006; Ravens-Sieberer et al., 2006). Consensus was reached regarding the conceptualisation and operationalisation of HRQoL, and physical, psychological and social aspects of health were retained as the broad domains in the assessment of HRQoL (Ravens-Sieberer et al., 2006). Current psychometric methods of classical and probabilistic test theory were used to determine the structure of the KIDSCREEN-52 and to fine-tune the scales (Embretson & Reise, 2000, as cited in Ravens-Sieberer et al., 2006). The KIDSCREEN-52 HRQoL index was examined to determine whether it met with the strict unidimensionality assumptions of the probabilistic Rasch model (Embretson & Reise, 2000, as cited in Ravens-Sieberer et al., 2006). However, since the KIDSCREEN-52 contains only items that have been verified to be generically age- and culture-relevant and comparable, the possibility exists that HRQoL aspects that are meaningful only for a particular subpopulation, may remain unconsidered (Ravens-Sieberer et al., 2006).

Numerous challenges emerge when using measures with different test populations. A key contentious issue is the impact of language on scores obtained. Measures are adapted from one language and culture to obtain a valid measurement to another cultural context (De Klerk, 2008). However, language difficulties such as the level of language comprehension (Matza, Swensen, Flood, Secnik & Leidy, 2004), low familiarity with item content (Ismail, 2010), and the literacy level of test-takers, can impact test scores. Administering a test in a language other than the mother tongue of the test-takers can also have a bearing on test scores.

The use of Western measures in non-Western cultures is particularly problematic since what constitutes QoL is to a large extent influenced by a person's beliefs and values, and is largely culturally determined (O'Connor, 2004). Cultures may vary not only in respect of the extent of subjective health, but also in the actual complaints expressed, and possibly in the exact meaning of a concept (Ravens-Sieberer et al., 2009). Culture-specific factors, such as the impact of philosophical traditions, linguistic influences and material living circumstances, have been found to influence how people perceive dimensions of well-being, and by extension, the construct itself, since QoL is directly derived from how individuals perceive the impact of experiences on their lives (Pflug, 2009). Hence, what is regarded as important to HRQoL in one country or culture may not be as important in other countries or cultures.

The impact of culture on testing a specific psychological construct must therefore be explored in order to adjust measurements so as to render them meaningful to the particular culture, as well as to obtain comparable or equivalent measures across cultures (De Klerk, 2008). An instrument may measure different constructs in different cultural groups, the relevant dimensions of the construct may

not be included in the formulation of item content, and the sampling of behaviours or characteristics associated with the construct may be inadequate for a particular cultural group. In addition, behaviours being tapped as indicators of a construct have the potential for differential interpretation (Van de Vijver & Tanzer, 2004). An exploration of the internal structure of a multi-dimensional instrument such as the KIDSCREEN-52, can elucidate the nature of the construct about which conclusions can be drawn from test-takers scores (Goodwin, 2000). This can be achieved by establishing the construct validity of the measure. 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). Hence, the statistical technique employed to assess whether the KIDSCREEN-52 measures the same underlying construct in a South African context, is factor analysis.

This entails the empirical assessment of the adequacy of a measure, and necessitates the establishment of validity and reliability. Reliability is an important feature of an assessment instrument because unreliability detracts from validity (Pesudovs, Burr, Harley & Elliot, 2007). Thus, if the measurements resulting from a test fluctuate drastically or are not stable over time, the test cannot be regarded as valid. Even though the KIDSCREEN-52 is a well-validated measure and its construct validity has already been established (Ravens-Sieberer et al., 2007), validity is not a conclusive feature of a test, but is relative to every specific purpose for which the test is used (Jooste, 2001). 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. Accordingly, the fact that the construct validity of the KIDSCREEN-52 has not been established in the South African context has informed the need for this study. In addition, a great paucity exists in South African literature regarding HRQoL assessment among children and adolescents, and validating existing measures for use in a South African context narrows this gap (see Taliep, 2010).

The overall aim of this study is therefore to assess the construct validity of the KIDSCREEN-52 in a South African context. The specific aims of this study are: (1) to explore the factor structure of KIDSCREEN-52 with a sample of South African adolescents, to see how it compares with the 10-dimensional structure identified by previous European studies, and (2) to assess the adequacy of the internal consistency estimates of each of the sub-scales of the KIDSCREEN-52, using Cronbach's alpha.

## **METHODS**

### **Participants**

Participants for the primary study were drawn from six public schools in 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 aim was to identify areas of greatest deprivation and high levels of crime, in order to reduce violence through the presidential urban renewal programme (Provincial Government of the Western Cape, 2005; Department of Provincial and Local Government, 2006).

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 criterion used in the choice of sample was that participants should be Grade 9 learners from six public schools within the Education Management and Development Centre (EMDC) South Metropole of the Western Cape Education Department. The Southern Metropole area was selected because of its accessibility as well as accounting for the highest number of nodal areas within the Western Cape (Savahl et al., 2009).

The 565 participants comprised 348 female and 218 male Grade 9 learners aged 14-18. The home language of more than half of the participants (52.9%) was English, but a significant percentage (39%) of participants did not have English as their first language. Of these, 25.5% spoke Xhosa and 13.8% spoke Afrikaans.

### **Procedure**

Ethical clearance was obtained from the University of the Western Cape Research Ethics Committee, the Western Cape Education Department and the school principals of the respective schools. Signed informed consent and assent documents were obtained from parents and learners respectively. The questionnaire was administered by the research team in the presence of staff members.

### **Measures**

The KIDSCREEN-52 is a self-report measure which is applicable to populations from 8 to 18 years of age (Ravens-Sieberer et al., 2005). The questionnaire assesses the frequency of behaviour/feelings or the intensity of an attitude, by using a 5-point Likert response scale with a recall period of one week. 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 into health services research and health reporting (Ravens-Sieberer et al., 2005). The KIDSCREEN-52 assesses 10 dimensions. Table 1 below delineates the dimensions and provides a brief description of the subscales, the number of items per dimension, and an example of items measured by each dimension.

**Table 1.** Description of subscales

	<b>Name of subscales* and example of items</b>	<b>Concept</b>
1.	Physical activities and health (5) e.g. Have you felt fit and well?	level of physical activity, energy and fitness.
2.	Feelings (6) e.g. Has your life been enjoyable?	psychological well-being including positive emotions and satisfaction with life.
3.	General mood (7) e.g. Have you felt that you do	depressive moods and emotions as well as worries and stressful feelings.

everything badly?

- |     |                                                                        |                                                                                             |
|-----|------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|
| 4.  | About yourself (5)<br>e.g. Have you been happy with the way you are?   | perception of self including whether appearance of body is viewed positively or negatively. |
| 5.  | Free time (5)<br>e.g. Have you had enough time for yourself?           | opportunity for create social and leisure time.                                             |
| 6.  | Family and home life (6)<br>e.g. Have your parent(s) understood you?   | relationship with parents and atmosphere at home.                                           |
| 7.  | Money matters (3)<br>e.g. Have you had enough money for your expenses? | financial resources.                                                                        |
| 8.  | Friends (6)<br>e.g. Have you spent time with your friends?             | relationships with peers.                                                                   |
| 9.  | School and learning (6)<br>e.g. Have you enjoyed going to school?      | perception of own capacity, comprising learning, concentration and feeling about school.    |
| 10. | Bullying (3)<br>e.g. Have you been afraid of other boys and girls?     | feeling rejected by peers.                                                                  |

---

\* Number of items in brackets

### **Statistical Analysis**

To examine the construct validity of the KIDSCREEN-52 in a South African context, this research explored its factor structure and the internal consistency reliability of the subscales. Data were analysed using the Statistical Program for the Social Sciences (SPSS, version 17.0) package. Since an accumulation of missing values owing to paired missing values may critically curtail the number of subjects on which the variance co-variance matrix is based (de Vet, Adèr, Terwee & Pouwer, 2005), the percentage of respondents with missing values was first calculated. The internal consistency reliability of the KIDSCREEN-52 was examined by means of Cronbach's alpha for the scale as a whole, as well as for the original subscales of the KIDSCREEN-52. The ensuing coefficient is rooted in the homogeneity of the items, with a high alpha emerging when items correlate well together (Hammond, 2006). In this way, reliability may be regarded as a manner of construct validation or of providing validity evidence (Hammond, 2006). It is generally agreed that the lower limit for Cronbach alpha is .70, although it may decrease to .60 in exploratory research (Hair, Black, Babin & Anderson, 2010).

To determine whether this study replicates the 10-dimensional structure identified by previous European and other studies, the factor structure of the KIDSCREEN-52 was examined by means of exploratory factor analysis (EFA). EFA is

deemed appropriate 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). In a critical assessment on the use of factor analysis, de Vet et al. (2005) have reviewed 13 journals to appraise the use of a health and well-being questionnaire. They regard the use of EFA, as opposed to confirmatory factor analysis, as 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 has not yet been used without a prior hypothesis. Since the present study aimed to ascertain whether the items of the KIDSCREEN-52 could be categorised into factors signifying the different dimensions of the construct HRQoL with a South African sample, it could be regarded as an appropriate statistical procedure to reveal the underlying structure of this instrument and to facilitate subsequent analysis.

As a first step, the factorability of the data was assessed through a visual inspection of the correlation matrix, and by means of the Kaiser-Meyer-Olkin Measure (KMO) of sampling adequacy, as well as Bartlett's test of sphericity. Factors were rotated obliquely using the direct oblimin rotation method. They were rotated to facilitate interpretation, but obliquely because of the assumption that the factors are theoretically related (Field, 2005; Hair et al., 2010). 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 given sample size (n=565) (Hair et al., 2010). The significance of a factor loading depends on the sample size, so a factor loading of .30 is significant for sample sizes of 350 or greater (Field, 2005; Hair et al., 2010). To determine the number of factors to retain, an *a priori* criterion of 10 factors was specified, since the researchers already knew how many factors to extract, and attempted to replicate previous studies and extract the same number of factors previously found in European studies (Hair et al., 2010). After a satisfactory factor solution was derived, based on a predetermined number of factors derived from research objectives and prior research, the final step entailed assigning meaning to the factors by careful interpretation of the pattern of factor loadings for a latent variable.

## **RESULTS**

### **Sample characteristics**

The final sample size (N = 565) of the current study fulfilled the criterion for factor analysis, with more than 10 participants per variable (Hair et al., 2010). The proportion of scale-level missing data (2.93%) was acceptable since it was less than the recommended maximum value of 25% (de Vet et al., 2005) and considered to be missing at random. Therefore, cases were excluded from the analyses listwise, in which case any participant with missing data for any variable was excluded (Field, 2005, p.646).

### Internal Consistency Reliability

The reliability analysis for the scale as a whole revealed a very reliable (.80) Cronbach's alpha coefficient. As shown in Table 1, the alpha coefficients for all the subscales ranged from .76 to .81, indicating satisfactory to good internal consistency.

**Table 2.** Cronbach's alphas for target sample

<b>Sub scales</b>	<b>Cronbach's alphas</b>
Physical activities and health	0.79
Feelings	0.76
General Mood	0.80
About yourself	0.79
Free time	0.76
Family and home life	0.77
Money matters	0.79
Friends	0.78
School and learning	0.80
Bullying	0.81

The distribution of the total well-being scores produced a bell-shaped curve, indicating that the data were normally distributed (see Taliep, 2010). A visual inspection of the correlation matrix revealed that all the variables correlated with one another, with the majority of correlations being significant ( $p < .05$ ). None of the items correlated very highly (.90) which diminished the concern for singularity in the data (see Taliep, 2010). There was therefore no need to consider eliminating any questions at this point of the analysis (Field, 2005). An investigation of the diagonal elements of the anti-image correlation matrices provided evidence that there were sufficient intercorrelations and common variance between the variables (.694 and above). This, together with the meritorious KMO Measure of Sampling Adequacy (.894) and the statistically significant Bartlett's Test of Sphericity ( $p < .001$ ) indicated that the current data were adequate for EFA. The *a priori* exploratory factor analysis extracted a 10-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 specified for extraction explained 56.91 % of the variance in the items. Most of the items loaded on the expected scales of the KIDSCREEN-52 except for the items in three of the scales ("Bullying", "About Yourself" and "Feelings") which loaded differently from the original questionnaire.



Feelings2		-	
		<b>606</b>	
Feelings3	.205	-	
		<b>582</b>	
Feelings1		-	-192
		<b>469</b>	
Feelings5		-	
		<b>682</b>	
Feelings4		-	
		<b>597</b>	
About Yourself1	.276	-	
		<b>355</b>	
Feelings6		<b>.311</b>	
		-	
		<b>315</b>	
About Yourself3			<b>.624</b>
About Yourself4			<b>.596</b>
About Yourself5			<b>.481</b>
Bullying1			.224
About Yourself2			.211

The first factor to emerge was “Parent relations and Home Life”, which comprised “Family 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 stable factor. Similar findings were observed with factors two, three, four, five, six and seven. These factors also had all the hypothesised items loading on them, and they all had three or more items loading with values higher than .50, which indicates stable factors. The eighth factor “Psychological Well-being” was also well defined, with three of the hypothesised items loading saliently on the factor (.606, .582, .469), indicating a stable factor.

While the last three factors 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. “Feelings” items 4 and 5 loaded on factor 9 “Self Perception” and “Feelings” item 6 cross-loaded significantly onto two different factors, namely 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)”, with all but one loading significantly. “Bullying” items were poorly presented in the factor solution. None of the three hypothesised bullying items had significant loadings; only one loaded insignificantly on factor 10, one loaded on moods and emotions, and one loaded on social support and peers. The communalities for the bullying items were also low (< .50), indicating that the “Bullying” dimension was not sufficiently represented by the factor solution.

**Table 4.** Factor names, items and loadings

Factor Number	Factor Name	Item number and item	Variable Loading
<b>Factor 1</b>	<b>Parent Relations and Home Life</b>	4. Have your parents had enough time for you?	.776
		6. Have you been able to talk to your parent(s) when you wanted to?	.764
		2. Have you felt loved by your parent(s)?	.725
<b>Factor 2</b>	<b>School Environment</b>	1. Have you been happy at school?	.748
		5. Have you enjoyed going to school?	.690
<b>Factor 3</b>	<b>Moods and Emotions</b>	2. Have you got on well at school?	.684
		3. Have you felt so bad that you didn't want to do anything?	.688
		4. Have you felt that everything in your life goes wrong?	.531
<b>Factor 4</b>	<b>Social Support and Peers</b>	2. Have you felt sad?	.512
		4. Have you and your friends helped each other?	.709
		5. Have you been able to talk about everything with your friends?	.635
<b>Factor 5</b>	<b>Physical Well-being</b>	6. Have you been able to rely on your friends?	.633
		4. Have you been able to run well?	.787
		3. Have you been physically active?	.708
<b>Factor 6</b>	<b>Financial Resources</b>	2. Have you felt fit and well?	.645
		1. Have you had enough money to do the same things as your friends?	.888
		3. Do you have enough money to do things with your friends?	.861
<b>Factor 7</b>	<b>Autonomy</b>	2. Have you had enough money for your expenses?	.822
		3. Have you had enough opportunity to be outside?	.586
		2. Have you been able to do things that you want to do in your free time?	.579
<b>Factor 8</b>	<b>Psychological Well-being</b>	4. Have you had enough time to meet friends?	.555
		2. Have you felt pleased that you are alive?	-.606
<b>Factor 9</b>	<b>Emotional Self</b>	3. Have you felt satisfied with your life?	-.582
		1. Has your life been enjoyable?	-.469
		5. Have you felt cheerful?	-.682
		4. Have you been in a good mood?	-.597

---

<b>Factor 10</b>	<b>perception Social Self Perception</b>	1. Have you been happy with the way you are?	-.355
		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

---

## DISCUSSION

This 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 discussion is a critical reflection on the importance of issues surrounding cross-cultural testing, test adaptation and the role of language in the assessment process. Various psychometric aspects were examined, including feasibility, factorability of the data, factorial structure, and reliability. The analysis of the internal structure of the KIDSCREEN-52 revealed 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.

Feasibility was assessed by examining the proportion of missing items in the KIDSCREEN-52 scales. Given that these values were minimal (2.93%) and considered to be missing at random, they were excluded from the analysis. Overall, the KIDSCREEN-52 demonstrated a strong internal consistency among a diverse low-income sample of South African adolescents in terms of both the total scale (.80) and individual domain scores (> .70). These high levels of  $\alpha$  engender confidence in the reliability of the KIDSCREEN-52 in a South African context. This therefore indicates that this measure is an internally reliable tool for the assessment of HRQoL in South African populations. These findings are similar to the  $\alpha$  coefficients reported for the KIDSCREEN-52 by the European KIDSCREEN group (.77 to .89) in all participating European countries, except for the case of one scale, namely “Bullying” (Social acceptance) which was found to be below .70 in one country (France) (Ravens-Sieberer et al., 2005). The findings of the current study are analogous with those of the Korean study conducted by Hong et al. (2006) with the translated K-KIDSCREEN-52. 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.

EFA 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, since the percentage of non-redundant residuals with absolute values >.05 was 5% (see Taliep, 2010). All the assumptions of factor analysis were also met. The factors expected to emerge from this analysis are: parent relations and home life, school

environment, moods and emotions, social support and peers, physical well-being, financial resources, autonomy, psychological well-being, self perception and social acceptance (bullying). While a ten-factor structure was extracted, the item loadings deviated slightly from the hypothesised loadings. The first seven factors all had three or more salient loadings with values higher than .50, indicating stable factors. Furthermore, all the items comprising these seven hypothesised dimensions loaded on the empirically derived factors as expected. The eighth factor (Psychological Well-being) had three of the hypothesised items loading significantly ( $> .40$ ), which rendered the factor valid. Thus, the original labels for these eight factors were retained. While the last two factors also had significant loadings, the hypothesised “Self perception” factor appears to have two factors, namely “Emotional Self-concept” and “Physical Self-concept”. The hypothesised “Bullying” factor, however, was poorly presented in the factor solution.

The substantial deviations from the hypothesised loadings and the one significant cross-loading of the last three questions with other domains can be explained in terms of the wording of questions and the close relationship between the various dimensions tapped by the questions. For example, “Feelings” 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 renamed “Emotional Self-perception”.

The loadings of these items may also be a reflection of how health and aspects of well-being were understood and expressed by respondents in South Africa. The last feelings item (“Have you had fun?”), for example, 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 have meant something totally different to the participants. Since having fun goes hand-in-hand with being with one’s friends, respondents could have interpreted the question as such. Also, the experience of the way influences among peers are encountered could be positive or negative, or with a positive or negative experience resulting from an individual’s self-concept. So, depending on the individual’s self-concept, his/her interaction with his/her peers will yield a corresponding response to the question “Have you had fun?” This indicates that aspects of well-being may be understood and expressed differently in different cultural contexts. Thus, behaviours being tapped by the construct have the potential for differential interpretation (Van de Vijver & Tanzer, 2004).

Factor 10 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. Perhaps it can be ascribed to the notion that the way we look and dress communicates to society who we are, or presents a desirable image of who we want them to think we are, and therefore be more socially acceptable. This

indicates that characteristics associated with the construct may not be adequate for a particular cultural group (Van de Vijver & Tanzer, 2004).

A further possible reason might be the fact that the KIDSCREEN-52 was standardised and normed on European English first-language-speaking samples, and that the first language of a significant percentage of the participants of the current study was not English (25.5% had Xhosa as a first language and 13.8% had Afrikaans as a first language). Since language difficulties can be an obstacle in assessment (Foxcroft & Roodt, 2005) questions may have been misunderstood. In addition, the level of language comprehension may have placed a lower limit on the appropriateness of certain questions to test-takers (Matza et al., 2004). The formulation of question-items, for example, might explain why the question regarding “Self-perception”, “Have you been worried about the way you look?” loaded on factor 10, which was named “Social acceptance” (Bullying) in the original scale. Baron, Byrne and Branscombe (2006) have shown that how others view one, impacts on how one views oneself. So this item could very well have been understood or interpreted as “Have you been worried about how others think you look?” and thus, loaded on the domain of social acceptance. Although the “Bullying” items did not load as expected, and did not have sufficient explanation, the “Bullying” scale displayed high internal consistency (.81), indicating that in an 11 factor solution these items may load consistently on an additional factor.

An individual’s grade level is also not an accurate reflection or indication of his/her reading ability or literacy level (Wasserman, Maja & Wright, 2010). 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. Language is regarded as the primary mediator in test performance, especially when the language in which a test is administered is not the first language of the test-taker (Foxcroft & Aston, 2006). This is especially true in a South African context, where English literacy levels and reading comprehension abilities are low among previously disadvantaged communities as well as Afrikaans speakers.

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 this study. The importance of establishing whether a test actually measures what it intends to measure cannot be overemphasised, particularly since results of HRQoL assessments are used in clinical decision-making, the evaluation of the quality of medical care, the estimation of healthcare needs, and an understanding of the outcomes of differences in health and well-being (Spieth & Harris, 1996).

Since this study is only the first step in the validation of the KIDSCREEN-52 in a South African context, it is suggested that future research should explore an 11 factor solution, and assess issues of bias and equivalence. In practice, statistical methods do not detect bias, but indicate that an item functions differently or provides different information for test-takers with the same ability, hence the term Differential Item Functioning (DIF) (Foxcroft & Roodt, 2005). To establish whether any bias exists, identified items should be further investigated to ascertain possible reasons for functioning differently (Foxcroft & Roodt, 2005). When DIF is investigated,

statistical procedures are used to compare test results of test-takers who belong to different cultural or language groups, but have the same ability, and therefore help monitor the validity and fairness of questionnaires (Foxcroft & Roodt, 2005; Lin & Rogers, 2005).

The same methods used to conduct an item analysis in proficiency tests can be used in non-proficiency measures (Kaplan & Saccuzzo, 2009). For example, as opposed to considering an item as right or wrong, it could be structured along lines of ascertaining whether it is or is not associated with a particular behavioural repertoire (Kaplan & Saccuzzo, 2009). Robitail et al.(2006) conducted a DIF analysis on the KIDSCREEN-27 (a shorter version of the KIDSCREEN-52) based on Item Response Theory modelling to determine whether items behaved in the same way in different countries, using Zumbo's logistic regression.

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 seven 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.

## **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), at present, 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 psychometric properties of the KIDSCREEN-52 with a sample of South African adolescents.

The results provide strong evidence that social self-perception and emotional self-perception are two different factors or a multi-dimensional dimension. It is also unclear whether the almost 40% of participants who were not English first-language speakers understood and interpreted the items as intended. Participants should have been tested in their first language instead, particularly because it becomes complex to unravel whether poor performance on the test is a result of language or communication difficulties or due to the fact that test-takers have a low level of understanding of the construct being assessed (Foxcroft, 2004). This initial psychometric analysis provided preliminary evidence to support the internal consistency reliability and validity of the KIDSCREEN-52 for sound measurement with some further suggested analysis and possible minor adaptation.

In line with Messick's (1989) notion of the consequential basis of construct validity, researchers should be cognisant of the fact that certain tests can have grave implications for individuals' lives. QoL instruments are increasingly used in the health sector to make important decisions about patients' health needs, and it is therefore vital that such instruments are reliable and valid for use in such populations. A test that aims to measure HRQoL, but in reality measures something

else, can lead to distorted interpretations of results and invalidates research using the instrument.

## References

- Anastasi, A., & Urbina, S. (1997). *Psychological Testing*. (7<sup>th</sup> ed.). Upper Saddle River, NJ: Prentice Hall International.
- Baron, R. A., Byrne, D., & Branscombe, N. R. (2006). *Social psychology*. (9<sup>th</sup> ed.). New York, NY: Pearson.
- 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). Retrieved from <http://pareonline.net/getvn.asp?v=10&cn=7>.
- DeCoster, J. (1998). *Overview of factor analysis*. Retrieved from <http://www.stat-help.com/notes.html>.
- De Klerk, G. (2008). Cross-cultural testing. In M. Born, C.D. Foxcroft & R. Butter (Eds.), *Online readings in testing and assessment*, International Test Commission. Retrieved from <http://www.intest.com.org/Publications/ORTA.php>.
- Department of Provincial and Local Government (2006). *Preliminary impact assessment for the Khayelitsha Mitchells Plain urban renewal programme*. Cape Town: MCA Africa. Retrieved from: <http://ebookbrowse.com/urp-socio-eco-profiling-of-nodes-khayelitsha-and-mitchells-plain-68200715561-pdf-d173224819>
- Detmar, S.B., Bruil, J., Ravens-Sieberer, U., Gosch, A., Bisegger, C., & The European KIDSCREEN Group. (2006). The use of focus groups in the development of the KIDSCREEN HRQOL questionnaire. *Quality of Life Research*, 15(8), 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. doi: 10.1007/s11136-004-5742-3
- Fallowfield, L (2009). *What is quality of life?* Retrieved from <http://www.whatisitseries.co.uk>.
- Field, A. (2005). *Discovering statistics using SPSS*. (2<sup>nd</sup> ed.). London: Sage Publications.
- Foxcroft, C.D. (2004). Planning a psychological test in a multicultural South African context. *South African Journal of Industrial Psychology*. 30(4), 8-15. Retrieved from <http://www.sajip.co.za/index.php/sajip/article/download/171/168>
- Foxcroft, C.D., & Aston, S. (2006). Critically examining language bias in the South African adaptation of the WAISS-III. *SA Journal of Industrial Psychology*, 32(4), 97-102. Retrieved from [www.sajip.co.za/index.php/sajip/article/download/243/240](http://www.sajip.co.za/index.php/sajip/article/download/243/240)
- Foxcroft, C., & Roodt, G. (2005). *An introduction to psychological assessment in the South African context*. (2<sup>nd</sup> ed.). New York, NY: Oxford University Press.

- Gill, T. M., & Feinstein, A. R. (1994). A critical appraisal of the quality of quality-of-life measurements. *Journal of the American Medical Association*, 272(8), 619–626. doi: 10.1001/jama
- Goldbeck, L., Schmitz, T. G., Besier, T., Herschbach, P., & Henrich, G. (2007). Life satisfaction decreases during adolescence. *Quality Life Research*, 16, 969–979. doi: 10.1007/s11136-007-9205-5
- 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. Retrieved from <http://www.info.gov.za/View/DownloadFileAction?id=70714>
- Hair, J. F., Black, W., Babin, B. J., & Anderson, R.E. (2010). *Multivariate data analysis*. (7<sup>th</sup> ed.). New Jersey, N.J.: Prentice Hall.
- Hammond, S. (2006). Using psychometric tests. In G.M. Breakwell, S. Hammond & C. Fife-Schaw (Eds.), *Research Methods in Psychology* (pp.182–209). London: Sage Publications.
- 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. *Journal of Korean Med Science*, 21(2), 295–299. doi: [10.3346/jkms.2006.21.2.295](https://doi.org/10.3346/jkms.2006.21.2.295)
- Hu, Y., Stewart-Brown, S., Twigg, L., & Weich, S. (2007). 'Can the 12-item General Health Questionnaire be used to measure positive mental health?'. *Psychological Medicine*, 37 (7), 1005 – 1013. doi: 10.1017/S0033291707009993
- Ismail, G. (2010). *Towards establishing the equivalence of the English version of the verbal analogies scale of the Woodcock Muñoz Language Survey (WMLS), across English and Xhosa first language speakers*. Unpublished master's thesis, University of the Western Cape, Cape Town, South Africa.
- Jooste, M. J. L. (2001). *Introduction to psychological testing*. Johannesburg: Rand Afrikaanse Universiteit.
- 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. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC10719/pdf/hsresearch00553-0150.pdf>.
- Kaplan, R.M., & Saccuzzo, D.P. (2009). *Psychological testing: Principles, applications, and issues*. (7<sup>th</sup> ed.). Belmont, CA.: Wadsworth.
- Lin, J., & Rogers, W. T. (2005). *Validity of the simultaneous approach to the development of equivalent achievement tests in English and French (Satge II)*. Alberta : Centre for Research in Applied Measurement and Evaluation, The University of Alberta. Retrieved from [www2.education.ualberta.ca/.../NCME%20paper%202005%20CRA](http://www2.education.ualberta.ca/.../NCME%20paper%202005%20CRA)
- 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. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/14720133>

- Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement* (3<sup>rd</sup> ed., pp.13-104). New York, NY: Macmillan Publishing Company.
- O'Connor, R. (2004). *Measuring quality of life in health*. London: Churchill Livingstone.
- O'Leary-Kelly, S. W., & Vokurka, R. J. (1998). The empirical assessment of construct validity. *Journal of Operations Management*, *16*, 387-405.
- 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.
- Provincial Government of the Western Cape. (2005). *The provincial social capital formation strategy with an emphasis on youth*. Cape Town: Provincial Government of the Western Cape. Retrieved from [http://www.westerncape.gov.za/Text/2006/3/provincial\\_social\\_capital\\_formation\\_strategy.pdf](http://www.westerncape.gov.za/Text/2006/3/provincial_social_capital_formation_strategy.pdf)
- 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. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9691720>
- Ravens-Sieberer, U., Erhart, M., Wille, N., Wetzels, 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., ...The European KIDSCREEN Group. (2005). The KIDSCREEN-52 quality of life measure for children and adolescent's development and first results from a European Survey. *Expert Review of Pharmacoeconomics & Outcomes Research*, *5*(3), 353-364.
- 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. *Psychosocial Medicine*, *4*, 1-17. Retrieved from <http://www.egms.de/en/journals/psm/2007-4/psm000040.shtml>.
- Ravens-Sieberer, U., Torsheim, T., Hetland, J., Volleberg, W., Cavallo, F., Jericek, H., ... 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. doi: 10.1007/s00038-009-5406-8
- Robitail, S., Simeoni, M., Erhart, M., Ravens-Sieberer, U., Bruil, J., Auquier, P., & 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. doi: 10.1016/j.jadohealth

- 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. *Health Services Research, 40*(5), 1584-1604.
- Singer, M.I., Anglin, T.M., Song, L., & Lunghofer, L. (1995). Adolescent's exposure to violence and associated symptoms of psychological trauma. *Journal of the American Medical Association, 273*, 477-482.
- Snyder, C.R., Hoza, B., Pelham, W.E., Rapoff, J., Ware, L., Danovsky, M., ... Stahl, K. (1997). The development and validation of the children's hope scale. *Journal of Paediatric Psychology, 22*, 399-421.
- 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.
- Taliep, N. (2010). *Evaluating the construct validity of the KIDSCREEN-52 quality of life questionnaire within a South African context utilising exploratory factor analysis: Initial validation*. Unpublished master's thesis, University of the Western Cape, Cape Town, South Africa.
- 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.
- Van de Vijver, F., & Tanzer, N.K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Psychological Application, 54*, 119-135. doi: 10.1016/j.erap
- 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). Retrieved from <http://www.hsag.co.za/index.php/HSAG/article/viewfile/469/488>.
- World Health Organisation. (1948). Constitution of the World Health Organisation: Basic document. Geneva, Switzerland, World Health Organisation.
- WHOQOL Group. (1994). Development of the WHO: Rationale and current status. *International Journal of Mental Health, 23*, 24-56. Retrieved from [http://www.who.int/mental\\_health/media/en/76.pdf](http://www.who.int/mental_health/media/en/76.pdf)