

Psychometric properties of the Self-Esteem Questionnaire for South African adolescents



Lauren G. Wild

Department of Psychiatry, University of Cape Town, Rondebosch, 7701, South Africa
e-mail: lwild@humanities.uct.ac.za

Alan J. Flisher

Department of Psychology and Mental Health, University of Cape Town, South Africa.

Arvin Bhana

Human Sciences Research Council, Durban, South Africa

Carl Lombard

Medical Research Council, Cape Town, South Africa

This article describes two studies investigating the reliability and factorial validity of scores on the Self-Esteem Questionnaire (SEQ) for assessing self-evaluations relating to peers, school, family, sports/athletics, body image and global self-worth in South African adolescents. Participants were 900 learners enrolled in Grades 8 and 11 at public schools in Cape Town, and 116 Grades 8 and 11 learners attending independent schools. The results provided general support for the six-factor structure proposed by DuBois, Felner, Brand, Phillips and Lease (1996) and indicated that SEQ scores have good internal consistency and adequate test-retest reliability for English-speaking South Africans. However, minor revisions are needed for all scale scores to have acceptable internal consistency when translated into isiXhosa or Afrikaans.

Keywords: adolescence; self-esteem; Self-Esteem Questionnaire (SEQ)

The past few decades have seen a resurgence of interest in the relationship between self-esteem and behaviour and affect. Low self-esteem (an average tone of self-disapproval or negative self-evaluation) has been implicated in many different child and adolescent problem behaviours (Jessor, 1991; McGee & Williams, 2000;

Vella, Persic & Lester, 1996). However, there are limitations to the instruments that are most commonly used to assess self-esteem.

Self-esteem has traditionally been viewed as a global, unidimensional construct. In recent years, however, there has been a growing interest in the potentially multidimensional structure of various self-system components. For example, there is extensive evidence that the phenomenology of the self-concept is far more adequately described by multidimensional models than by global, unidimensional measures, and multidimensional instruments such as the Self-Perception Profile for Children/Adolescents (Harter, 1985, 1988) are among the most psychometrically sound measures available for assessing the self-concept. However, DuBois, Felner, Brand, Phillips and Lease (1996) point out that while such measures assess perceived competencies in various domains, they do not focus on the degree of satisfaction with such competencies, as would be the case in assessing self-esteem. Moreover, most of the measures traditionally used to measure self-esteem have questionable content validity, and include many items that do not have a self-evaluative component and therefore may not be assessing self-esteem *per se* (DuBois et al., 1996). The widely used Coopersmith (1967) Self-Esteem Inventory, for example, includes items that ask for non-evaluative self-descriptions (e.g., 'I would rather play with children younger than I am') and perceptions of others' views of the self (e.g., 'My parents understand me').

DuBois et al. (1996) have therefore developed a new instrument, the Self-Esteem Questionnaire (SEQ), for assessing multiple domains of self-esteem in young adolescents. The SEQ comprises six sub-scales assessing evaluations of the self relating to five salient contexts or domains of experience for this age group (peers, school, family, sports/athletics and body image) as well as global feelings of self-worth. DuBois et al. (1996) explored the psychometric properties of their scale using a demographically diverse (but predominantly white) community sample of 1 800 adolescents enrolled in Grades 5 to 8 in a Midwestern American state. They found that scores obtained on the instrument had good internal consistency reliability (coefficient alphas for each subscale ranging from 0.81 to 0.91), and provided evidence for the factorial validity of the proposed sub-scale structure. Subsequent analyses conducted with a smaller sample of 225 adolescents attending Grades 5 to 9 in a single junior high school located in a Midwestern community provided support for the SEQ's convergent and discriminant validity across self-report, interview and parent-report forms. The authors also reported that ratings of specific dimensions of self-esteem exhibited differentially strong associations with measures of contextual experiences (perceived support, daily stressors and major events) in corresponding areas. In addition, test-retest reliability was calculated over a period of two weeks for a sub-sample of 52 learners and found to be adequate (*rs* ranging from 0.74 to 0.84).

These studies provide preliminary evidence for the reliability and validity of scores obtained on the SEQ. However, the scale has not been used previously in

South Africa, and its psychometric properties for samples outside the United States of America (US) are unknown. DuBois et al.'s (1996) findings regarding the reliability of scores on the instrument cannot be generalised to a South African population as it differs from an American population in terms of factors such as language, literacy and cultural expression, which might affect the understanding of questions, and hence their reliability (Shrout, 1995). Concepts and terms describing subjective emotional experiences and states are not easy to translate into different languages, and there are likely to be local variations in semantics even between different English-speaking populations (Robertson & Berger, 1994).

This article describes two studies designed to investigate the reliability and factor validity of scores on the SEQ as a multidimensional measure of self-esteem in adolescents in Cape Town. Study 1 investigated the internal consistency of the SEQ scale scores amongst a representative sample of English-, Afrikaans- and isiXhosa-speaking adolescents, and used confirmatory factor analyses to examine the hypothesised subscale structure of the measure. Study 2 investigated the test-retest reliability of scores obtained on the questionnaire with a smaller sample of English-speaking adolescents.

STUDY 1

METHOD

Sample

The study population consisted of all learners in Grades 8 and 11 attending public schools in Cape Town. Schools were stratified by postal code groupings, as areas with the same postal code tend to be relatively homogeneous with regard to socio-economic status and racial composition. Thirty-nine schools were selected such that the proportion of the selected schools in a particular stratum was directly proportional to the number of learners in that stratum. Within each stratum, the probability of selecting a school was proportional to the number of learners in that school. Two classes were randomly selected from each participating grade, and 40 learners were randomly selected from their combined class lists. An additional 5 learners were selected as replacements for a maximum of 5 absent learners. This procedure yielded a total sample of 2 946 learners, of whom one-third were randomly selected to complete the self-esteem measure. The remaining learners completed other questionnaires that are not included in this article.

Of the participating learners ($N = 900$), 505 (56%) were female. The mean age of the sample was 15.6 years (range = 12-26, $SD = 2.16$). Of the learners, 50% ($n = 451$) described themselves as coloured (derived from Asian, European and African ancestry), 24% ($n = 220$) as black, 21% ($n = 197$) as white and three learners described themselves as Asian. The remaining 29 learners did not report their racial

classification. The original English-language version of the questionnaire was completed by 392 learners (44%); 371 (41%) completed the Afrikaans-language version and 137 (15%) completed the isiXhosa language version.

Instrument

Adolescents' self-esteem was assessed using the multidimensional SEQ developed by DuBois et al. (1996). The measure consists of 42 items, each of which is rated on a 4-point scale ranging from *strongly disagree* to *strongly agree*. Each item is scored 1 to 4, with higher values indicating higher self-esteem. In order to guard against possible bias associated with response style, ten items are negatively worded and reverse scored. The SEQ is divided into six sub-scales. One of these comprises items directly assessing adolescents' global self-worth (e.g., 'I am happy with myself as a person'). The remaining sub-scales assess adolescents' self-evaluations with respect to five salient contexts or domains of adolescents' experience: peers (e.g., 'I am as popular with kids my own age as I want to be'), school (e.g., 'I am good enough at math'), family (e.g., 'I am happy with how much my family loves me'), sports/athletics (e.g., 'I am as good at sports/physical activities as I want to be') and body image (e.g., 'I like my body just the way it is').

The SEQ was translated from English into isiXhosa and Afrikaans, the other languages most commonly spoken in Cape Town. The translated versions of the questionnaire were back translated into English by people whose home language was either isiXhosa or Afrikaans. These back-translated versions were compared with the original version, and any discrepancies were resolved by negotiation between the original translators and those doing the back translations. Translated versions of the questionnaire were then piloted in small groups and classrooms in order to ensure that the translations were adequate.

Procedure

Permission to conduct the study was obtained from the Western Cape Education Department and the principals of the selected schools. The questionnaires were administered in the classroom by members of the research team, and participants were assured that their responses would be anonymous and confidential. No members of the school staff were present during the administration of the questionnaires, and care was taken to ensure that the learners were seated such that they could not see the responses of their classmates. Learners were informed that they could choose not to participate in the study as a whole or to omit certain questions.

RESULTS

Descriptive statistics

Descriptive statistics for the self-esteem scales are presented in Table 1. Means for the majority (90%) of individual items were above the scale midpoint of 2.5. The peers, family, body image and global self-esteem scales had means above the midpoint of their possible ranges, and response distributions for these scales and for most items showed slight to moderate degrees of negative skew, reflecting a tendency toward positive self-evaluation. Responses on the school and sports self-esteem scales were more normally distributed. The range of scores on each scale was sufficiently broad to include scores at least 1.5 *SDs* above the sample mean. Thus, there was substantial variability in item response distributions for each scale, and there were no notable 'floor' or 'ceiling' effects.

Table 1. Descriptive statistics for Self-Esteem Questionnaire scale scores

Self-Esteem Scale	<i>M</i>	<i>SD</i>	Range	Skewness
Peers (8 items)	22.44	4.36	8-32	-0.35
School (8 items)	20.79	5.12	8-32	-0.05
Family (8 items)	23.96	5.00	8-32	-0.43
Body image (4 items)	10.97	2.90	4-16	-0.28
Sports/athletics (6 items)	15.20	3.52	6-24	-0.05
Global self-esteem (8 items)	22.46	4.39	8-32	-0.12

Internal consistency

Measures of internal consistency for scores on each scale are presented in Table 2. Internal consistency (reliability) of scale scores was assessed using Cronbach alpha. For the sample as a whole, scores on all of the self-esteem scales, except sports/athletics ($\alpha = 0.67$), were found to have internal consistency coefficients above 0.70, the recommended minimum level for research use (Nunnally & Bernstein, 1994). As can be seen in Table 2, alpha coefficients for scores on the other scales ranged from 0.74 for the body image scale to 0.85 for the family self-esteem scale. Deleting one item (Item 28) (item total correlation = 0.06; factor loading = 0.04) from the sports self-esteem scale raised the alpha coefficient for this scale to an acceptable 0.74.

Alpha coefficients were then examined separately for each language group (English, Afrikaans and isiXhosa). Scores on the peer, school and family self-esteem scales were found to have acceptable internal consistency ($\alpha \geq 0.70$) for all language versions, with alpha coefficients for these three sub-scales ranging from 0.74 to 0.88.

In addition, scores on the Afrikaans translation of the body-image and global self-esteem scales had acceptable internal consistencies. However, scores on the isiXhosa version of the body image and global self-esteem sub-scales had unacceptable internal consistencies, and scores on the sports/athletics self-esteem scale had a low internal consistency in both the isiXhosa and Afrikaans versions. Scores on the body image scale had a coefficient alpha of only 0.47 for the isiXhosa version, which rose to 0.78 if one item (Item 22) (item total correlation = -0.25) was deleted. Deleting one item (Item 42) (item total correlation = -0.43) from the global self-esteem scale raised the alpha coefficient for scores on the isiXhosa version of the questionnaire from 0.47 to 0.65. Similarly, deleting one item (Item 11) (item total correlation = 0.17 for the Afrikaans version, and -0.50 for the isiXhosa version) from the sports/athletics scale increased alpha from 0.68 to 0.72 for scores on the Afrikaans questionnaire, and from 0.13 to 0.52 for scores on the isiXhosa questionnaire. Deleting one further item (Item 28) (item total correlation = -0.51) from the isiXhosa questionnaire raised the alpha coefficient for scores on the remaining 4 items to 0.83.

Table 2. Internal consistency estimates for Self-Esteem Questionnaire scale scores

Self-Esteem Scale	Cronbach α			
	Total	English	Afrikaans	isiXhosa
Peers (8 items)	0.78	0.83	0.74	0.76
School (8 items)	0.84	0.86	0.83	0.79
Family (8 items)	0.85	0.88	0.85	0.74
Body image (4 items)	0.74	0.78	0.77	0.47
Sports/athletics (6 items)	0.67	0.75	0.68	0.13
Global self-esteem (8 items)	0.76	0.83	0.74	0.47

Notes: Missing data were deleted casewise. Valid ns range from 764 to 821 for the entire sample; from 350 to 368 for the English sub-sample; from 312 to 339 for the Afrikaans sub-sample; and from 98 to 114 for the isiXhosa sub-sample.

Factor structure

Confirmatory factor analyses (CFAs) were used to investigate the goodness-of-fit of DuBois et al.’s (1996) six-factor model to the data obtained in this study. All 42 SEQ items were included in these analyses, which were performed using the SEPATH module of the Statistica computer programme (StatSoft, 1984). The CFA model (a) fixed all factor variances at unity, (b) allowed all factors to correlate freely, (c) freed all hypothesised factor loadings and constrained others at zero, and (d) allowed for correlated measurement error (i.e., correlations between residual variables) for the ten reverse-scored items as a result of method of assessment. The factor structure of this model was ‘pure’ in that each observed variable (SEQ item) was allowed to load on only one factor.

The Bentler Comparative Fit Index (CFI), the Bentler-Bonett Non-Normed Fit Index (NNFI) and the Joreskog GFI were selected as the most appropriate fit indices for this study, based on the recommendations of Hu and Bentler (1995), Gerbing and Anderson (1992) and Marsh et al. (1988). Fit indices obtained for this model were as follows: CFI = 0.90; NNFI = 0.88; GFI = 0.85. Two of these indices are lower than the minimum acceptable level of 0.90 proposed by Bentler and Bonett (1980) as a rule of thumb in using confirmatory factor analysis to evaluate the adequacy of factor models. However, the RMS standardised residual for the model was only 0.07, and the largest standardised residual (i.e., discrepancy between observed and reproduced correlations) was 0.19. This suggests that the model explains the correlations quite well overall, and is only slightly wrong for some variables (Hu & Bentler, 1995). All but one of the SEQ items had statistically significant loadings on their hypothesised factors; the exception was Item 28 ('I wish it were easier for me to learn new sports/physical activities'), which had a factor loading of only 0.04 ($p = 0.26$) on the sports/athletics scale. Deleting this item had a negligible effect on the goodness-of-fit indicators.

The CFA results also indicated, however, that correlations between the factors were generally quite high, ranging from 0.57 for the correlation between the family and school factors to 0.90 for the body image and global self-esteem factors (all $ps < 0.001$). Values of r^2 ranged from 0.32 to 0.81, indicating that between 32% and 81% of the variability in each factor is directly predictable from the variability in another factor. This raises the possibility that a single self-esteem factor might fit the data as well as, or better than, the six-factor model. DuBois et al.'s (1996) six-factor model was therefore compared with a single-factor model. Results indicated that the six-factor model, statistically, provided a significantly better fit to the data than a model in which all items were hypothesised to load on a single self-esteem factor ($\Delta\chi^2(87) = 1\,506.17, p < 0.001$).

STUDY 2

METHOD

Sample

The research participants comprised 116 learners enrolled in Grade 8 ($n = 68$) and Grade 11 ($n = 48$) at independent (private) secondary schools in Cape Town, South Africa. It was not possible to conduct the study at public-sector schools, since access was denied because of the disruption that would have been caused by administering questionnaires on two occasions. However, the independent schools had a mainstream educational philosophy similar to that of public-sector schools and did not strongly identify with any specific religious orientation. Four out of the five schools approached agreed to participate. One school was for boys only, one for girls only, and two were co-educational. All four schools admitted learners of all races and used English as the medium of instruction.

Of the 358 learners in the participating classes, one third ($n = 116$) were randomly selected to complete the SEQ. The remaining learners completed other questionnaires that are not included in this article.

Of the participants, 71 (61%) were male. Their mean age was 14.6 years (range = 12-18, $SD = 1.60$). Most of the learners (77%) were white, 18 (15%) identified themselves as coloured, 6 (5%) as black, and 3 as Asian. The vast majority (97%) reported that English was one of the languages that they spoke at home.

Procedure

Permission to carry out the study was obtained from the principals of the selected schools. The original English-language version of the SEQ was administered in the classroom by members of the research team, following the procedures described for Study 1. As in Study 1, learners were informed that participation was voluntary and that they could choose to omit selected questions.

The questionnaires were administered on two occasions, 10 to 14 days apart. This period represents a compromise between the possibility that self-esteem could change and the ability of the respondents to recall their responses from the first administration when completing the questionnaire on the second occasion (Brenner et al., 1995). The following procedure was carried out to enable matching of the questionnaires that were completed at Time 1 (T1) and Time 2 (T2), while keeping the questionnaires anonymous. At the end of the first session the learners were requested to write the identifying number of their questionnaire on a piece of paper supplied by the researchers, put this piece of paper in an envelope, seal the envelope, sign across the seal and write their name on the outside of the envelope. This sealed envelope was then returned to the learner when the second administration session took place. The learners opened the envelopes and transferred the questionnaire number to the new questionnaire.

RESULTS

Test-retest reliability for scores on the individual items comprising the SEQ was computed by two methods: observed agreement and Cohen's weighted kappa (κ_w). The former refers to the percentage of cases where there was agreement between the two administrations and the latter indicates whether agreement between the two administrations is beyond that which would be expected by chance (Cohen, 1960). Weighted kappa was used in preference to simple kappa because the SEQ items have four ordered response categories (ranging from *strongly disagree* to *strongly agree*). Weighted kappa quantifies the relative seriousness or extent of each possible disagreement by assigning maximal weights for exact agreement and minimal weights for extreme disagreement (Fleiss, 1981; Landis & Koch, 1977; Maclure & Willett, 1987). The results of these analyses are displayed in Table 3.

The observed agreement for scores on each item ranged from 0.53 to 0.79. As kappa is adversely affected by inconsistencies in the distributions of the observed response profiles on the two occasions the questionnaire was administered, McNemar's test was used to examine whether the marginal proportions were homogeneous. If the p -value for McNemar's test is < 0.05 , this implies that there is a significant difference between the prevalence rates for the two administrations and the calculation of kappa is not appropriate (Landis & Koch, 1977; Maclure & Willett, 1987). No significant McNemar's statistics were found, therefore, kappa was calculated for all items.

The descriptive terms of Landis and Koch (1977) are commonly used to characterise values of both simple and weighted kappa: < 0.00 , poor; $0.00 - 0.20$, slight; $0.21 - 0.40$, fair; $0.41 - 0.60$, moderate; $0.61 - 0.80$, substantial; and $0.81 - 1.00$, almost perfect. According to these conventions, none of the SEQ item scores evidenced poor or slight agreement. Weighted kappa values indicated fair agreement in 6 instances, moderate agreement for scores on 33 items, and substantial agreement for 3 items. Since it was thought desirable to maintain the integrity of the scales as much as possible, all items were therefore retained and summed to form the six SEQ scales.

Test-retest reliability of scores on each of these scales was calculated using Pearson product-moment correlation coefficients. Missing data were deleted listwise for these analyses. Correlation coefficients ranged from 0.73 for the sports/athletics self-esteem scale to 0.83 for the school and body image scales (all $ps < 0.001$). Values of r^2 ranged from 0.53 to 0.69, indicating that more than half of the variability in scores on each self-esteem scale at the second administration of the questionnaire was directly predictable from the variability in scores on that scale at the first administration. The mean r for the six sub-scales was 0.79 ($r^2 = 0.62$).

DISCUSSION

This study was designed to investigate the factorial validity, internal consistency and test-retest reliability of scores on the SEQ for a South African population. Overall, the results of this study support the use of most of the items and sub-scales of the SEQ for assessing multiple dimensions of self-esteem amongst South African adolescents in high schools.

Consistent with DuBois et al.'s (1996) findings with American adolescents, the descriptive statistics indicated a slight tendency towards positive self-evaluation on the SEQ. Nevertheless, there was a wide range of responses to all the self-esteem items.

In general, the results of the confirmatory factor analysis supported the six-factor structure proposed by DuBois et al. (1996). Only one item had a statistically non-significant factor loading on the relevant (sports/athletics) self-esteem subscale. In addition, the results confirmed that DuBois et al.'s six-factor model provided a better

Table 3. Test–retest reliability for Self-Esteem Questionnaire items

Item	N	Observed agreement (%)	Weighted kappa
1	112	65	0.43
2	111	55	0.33
3	111	70	0.60
4	108	68	0.57
5	109	63	0.51
6	108	70	0.45
7	107	60	0.46
8	110	66	0.56
9	109	67	0.53
10	109	67	0.59
11	111	66	0.47
12	110	61	0.53
13	110	59	0.45
14	109	66	0.62
15	108	62	0.46
16	105	64	0.51
17	109	57	0.30
18	107	72	0.47
19	109	61	0.35
20	110	62	0.49
21	111	69	0.48
22	109	68	0.60
23	110	57	0.33
24	109	72	0.55
25	110	71	0.36
26	106	71	0.57
27	109	72	0.63
28	110	61	0.44
29	109	53	0.39
30	111	60	0.45
31	110	69	0.52
32	110	65	0.55
33	110	63	0.42
34	106	74	0.60

Table 3. (Continued)

Item	N	Observed agreement (%)	Weighted kappa
35	107	78	0.55
36	106	79	0.67
37	109	69	0.58
38	105	70	0.50
39	108	71	0.48
40	105	64	0.58
41	110	71	0.54
42	105	64	0.57

fit to the data than a model in which all the items were hypothesised to load on a single self-esteem factor.

Also consistent with DuBois et al.'s (1996) findings, the highest alpha coefficients were obtained for scores on the family self-esteem scale, and the lowest for scores on the sports/athletics self-esteem scale. However, levels of internal consistency obtained in this study (coefficient alphas ranging from 0.67 to 0.85 for scores on each scale) were lower than those reported by DuBois et al. (1996) (0.81 to 0.91), albeit still mostly acceptable. An examination of internal consistency coefficients for each language version of the questionnaire indicated that there were seven items for which scores tended to have relatively low item-total correlations as well as relatively low factor loadings (i.e., ≤ 0.36) across language groups. These were all negatively worded, reverse-scored items. Thus, the fact that scores on 7 of the 10 reverse-scored items in the questionnaire had relatively low item-total correlations suggests that some participants may have answered these items erroneously because they were failing to read the questions carefully due to fatigue or laziness, and were simply answering according to a response set.

Analyses also indicated, however, that internal consistency was highest for scores on the English version of the questionnaire, slightly lower for the Afrikaans version, and notably lowest for the isiXhosa version. This suggests that the translations of some items might be inadequate. It may be that certain English concepts and phrases cannot be meaningfully translated into isiXhosa, a Bantu language that has a very different grammatical structure from the West Germanic languages of English and Afrikaans. However, isiXhosa-speaking research assistants have suggested an alternative (improved) translation for the phrase, 'I wish', which appears in four of the seven items with low internal consistencies (e.g., 'I wish my friends liked me more than they do'). Revising these items might therefore help to improve the internal consistency of scores on the isiXhosa translation of the SEQ.

Thus, overall it appears that scores on the English version of the SEQ have good internal consistency for a South African population. In addition, scores on all of the self-esteem sub-scales, except sports/athletics, had good internal consistency when these scales were translated into Afrikaans. However, certain items in the sports/athletics, body image and global self-esteem sub-scales may have to be revised or omitted for these scales to have acceptable internal consistencies when translated into isiXhosa. In addition, minor alterations may have to be made to the Afrikaans version of the sports/athletics subscale.

The results of Study 2 also provided evidence in favour of the test–retest reliability of scores on the SEQ items and sub-scales. The weighted kappa values obtained indicated that test-retest reliabilities for responses to the individual SEQ items ranged from fair to substantial, with scores on most items evidencing moderate agreement across the two administrations of the questionnaire. Test–retest reliability coefficients for scores on the six self-esteem sub-scales ranged from 0.73 for the sports/athletics subscale to 0.83 for the school and body image sub-scales. These are very similar to the two-week test–retest reliability coefficients obtained by DuBois et al. (1996), who reported that *rs* ranged from 0.74 for the sports/athletics subscale to 0.84 for the school subscale. The mean *r* for the six sub-scales was 0.79, which is the same as the mean of the *rs* obtained by DuBois et al. (1996), and similar to test–retest reliability coefficients obtained for established self-esteem scales such as the Rosenberg (1979) Self-Esteem Scale (Bosson et al., 2000) and the Coopersmith (1967) Self-Esteem Inventory (Bedeian et al., 1977; Ryden, 1978) over a period of four to six weeks. Self-esteem measures are unlikely to yield very high or almost perfect test–retest reliability coefficients. There is considerable evidence that situational factors can lead to short-term changes in self-evaluation (Harter, 1998), and some actual changes in feelings of self-esteem are therefore likely to occur between two administrations of a questionnaire. Thus, these findings suggest that SEQ scores have adequate test-retest reliability amongst an English-speaking, South African sample.

In interpreting these results, it is important to remember that this research was conducted in the Cape Town metropolitan area, and generalisations to adolescents in rural areas and/or other parts of the country must be made with appropriate caution. Nevertheless, the findings of these studies have provided preliminary evidence that scores on the SEQ are a reliable measure of multiple dimensions of self-esteem for South African adolescents, although certain minor alterations may need to be made to the existing isiXhosa and Afrikaans translations of some of the sub-scales.

ACKNOWLEDGEMENTS

Financial support for this research was provided by the World Health Organisation (WHO) Programme on Substance Abuse, the United Nations Development Programme, the South African Medical Research Council and the Medical Faculty Research Committee of the University of Cape Town. The authors would like to

thank the Western Cape Education Department, the principals, staff and learners of the schools that participated in the study, Janet Evans and Lisa Wegner for assistance with the fieldwork, and Martie Muller for her contributions to data management.

REFERENCES

- Bedeian, A. G., Teague, R. J. Jr. and Zmud, R. W. (1977). Test–retest reliability and internal consistency of short-form of Coopersmith’s self-esteem inventory. *Psychological Reports*, *41*, 1041–1042.
- Bentler, P. M. and Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588–606.
- Bosson, J. K., Swann, W. B. Jr. and Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology*, *79*, 631–643.
- Brener, N. D., Collins, J. L., Kann, L., Warren, C. W. and Williams, B. I. (1995). Reliability of the Youth Risk Behavior Questionnaire. *American Journal of Epidemiology*, *141*, 575–580.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.
- Coopersmith, S. (1967). *The antecedents of self-esteem*. San Francisco, CA: Freeman.
- DuBois, D., Felner, R. D., Brand, S., Phillips, R. S. C. and Lease, A. M. (1996). Early adolescent self-esteem: A developmental-ecological framework and assessment strategy. *Journal of Research on Adolescence*, *6*, 543–579.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: Wiley.
- Gerbing, D. W. and Anderson, J. C. (1992). Monte Carlo evaluations of goodness-of-fit indices for structural equation models. *Sociological Methods and Research*, *21*, 132–160.
- Harter, S. (1985). *Manual for the self-perception profile for adolescents*. Denver, CO: University of Denver.
- . (1988). *Manual for the self-perception profile for children*. Denver, CO: University of Denver.
- . (1998). The development of self-representations. In N. Eisenberg (Ed.), *Handbook of child psychology (5th Ed.): Vol. 3. Social, emotional and personality development* (pp. 553–617). New York: John Wiley & Sons.
- Hu, L.-T. and Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues and applications* (pp. 76–99). Thousand Oaks, CA: Sage.
- Jessor, R. (1991). Risk behavior in adolescence: A psychosocial framework for understanding and action. *Journal of Adolescent Health*, *12*, 597–605.
- Landis, J. R. and Koch, G. C. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174.
- Maclure, M. and Willett, W. C. (1987). Misinterpretation and misuse of the kappa statistic. *American Journal of Epidemiology*, *126*, 161–169.
- Marsh, H. W., Balla, J. R. and McDonald, R. P. (1988). Goodness-of-fit indices in confirmatory

- factor analysis: The effect of sample size. *Psychological Bulletin*, 103, 391–410.
- McGee, R. and Williams, S. (2000). Does low self-esteem predict health compromising behaviors among adolescents? *Journal of Adolescence*, 23, 569–582.
- Nunnally, J. C. and Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Robertson, B. and Berger, S. (1994). Child psychopathology in South Africa. In A. Dawes & D. Donald (Eds), *Childhood and adversity: Psychological perspectives from South African research* (pp. 154–176). Cape Town, South Africa: David Philip.
- Rosenberg, M. (1979). *Conceiving the self*. New York: Basic Books.
- Ryden, M. B. (1978). An adult version of the Coopersmith Self-Esteem Inventory: Test-retest reliability and social desirability. *Psychological Reports*, 43, 1189–1190.
- Shrout, P. E. (1995). Reliability. In M. T. Tsuang, M. Tohen and G. E. P. Zahner (Eds), *Textbook in psychiatric epidemiology* (pp. 213–227). New York: Wiley-Liss.
- StatSoft (1984). *Statistica (Version 6)* [Computer software]. Tulsa, OK: Author.
- Vella, M. L., Persic, S. and Lester, D. (1996). Does self-esteem predict suicidality after controls for depression? *Psychological Reports*, 79, 1178.