

Psychometric properties of the Chinese (Cantonese) versions of the KIDSCREEN health-related quality of life questionnaire

Johan Y. Y. Ng¹ · Angus Burnett² · Amy S. Ha¹ · Kim Wai Sum¹

Accepted: 23 March 2015 / Published online: 11 April 2015
© Springer International Publishing Switzerland 2015

Abstract

Purpose The KIDSCREEN questionnaire assesses health-related quality of life in children and adolescents. In this study, the questionnaire was translated into Chinese (Cantonese) and administered to two independent groups of Hong Kong Chinese students. Various aspects of reliability and validity of the KIDSCREEN-52 and KIDSCREEN-27 questionnaires were examined.

Methods The translated KIDSCREEN-52 questionnaire (and thus the KIDSCREEN-27 as a subset of the long form) was administered to cohorts of 1379 and 555 students. Confirmatory factor analysis and partial credit Rasch analyses were used to examine the underlying structure of the instrument. Test–retest reliability, convergent, and divergent validity were also examined.

Results The 10-factor structure of the Chinese KIDSCREEN-52 was supported. However, the original five-factor model of the KIDSCREEN-27 was not supported as seven distinct factors were found. Internal consistency and test–retest reliability were acceptable. Evidence supporting convergent and divergent validity was found.

Conclusions Results from the two studies supported the use of the translated Chinese version of KIDSCREEN questionnaire children and adolescents' health-related quality of life. Further research is required to examine possible cultural or language differences between the original version and the translated Chinese version of the questionnaire.

Keywords Confirmatory factor analysis · Partial credit Rasch model · Validity · Reliability

Introduction

Health-related quality of life (HRQoL) is considered as an important well-being outcome by clinical practitioners and researchers [1], and the KIDSCREEN questionnaire [2, 3] is a validated measure of the construct for children and adolescents. The 52-item questionnaire (“KIDSCREEN-52”) measures ten facets (physical well-being, psychological well-being, etc. as per Table 1) of HRQoL, while a shortened, 27-item, version (“KIDSCREEN-27” with five subscales) has been obtained from the longer version of the questionnaire for increased clinical utility (full detail of the questionnaires is provided elsewhere [2]). These questionnaires have been translated into multiple languages [4–6]; however, to the best of our knowledge, there is no existing version for use in Hong Kong. In the current study, KIDSCREEN items were translated into Chinese (specifically Cantonese). The translated version of the full questionnaire was then administered to two cohorts of Hong Kong Chinese students. Aspects of reliability and validity of the KIDSCREEN-52 and KIDSCREEN-27 questionnaires were subsequently examined.

Methods

Participants

Two cohorts including 1379 (12.4 ± 1.9 years; 47.4 % female) and 555 (11.7 ± 2.2 years; 51.0 % female; two administrations 2 weeks apart with 95 % retention)

✉ Amy S. Ha
sauchingha@cuhk.edu.hk

¹ Department of Sports Science and Physical Education, The Chinese University of Hong Kong, Shatin, Hong Kong

² ASPETAR, Qatar Orthopaedic and Sports Medicine Hospital, Doha, Qatar

Table 1 Psychometric properties of the KIDSCREEN-52 and KIDSCREEN-27 subscale scores in Cohort 1

Subscale	Number of items	Mean	SD	Cronbach's alpha	Missing (%)	Ceiling (%)	Floor (%)	Range of factor loadings in CFA
KIDSCREEN-52								
Physical Well-being ^a	5	3.63	.76	.84	.1	5.9	.0	.67–.92
Psychological Well-being	6	3.70	.86	.95	.0	12.4	.8	.88–.94
Moods and Emotions	7	3.80	.78	.91	.5	7.8	.4	.76–.90
Self Perception	5	3.74	.73	.74	.4	7.0	.1	.31–.92
Autonomy	5	3.38	.88	.89	.5	8.0	.6	.84–.89
Parent Relation and Home Life	6	3.60	.87	.93	1.2	9.9	.2	.79–.95
Financial Resources	3	3.30	1.04	.90	1.3	11.8	3.8	.85–.94
Peers and Social Support	6	3.68	.83	.92	1.7	10.9	.5	.79–.92
School Environment	6	3.59	.82	.92	.2	7.5	.3	.79–.94
Bullying	3	4.20	.83	.87	1.7	30.1	1.5	.86–.89
KIDSCREEN-27								
Physical Well-being ^a	5	3.63	.76	.84	.1	5.9	.0	.63–.91
Psychological Well-being	4	3.62	.84	.91	.0	10.0	.6	.82–.92
Moods and Emotions	3	3.88	.81	.84	.6	13.2	.7	.81–.88
Autonomy and Parent Relation	5	3.47	.84	.87	.1	7.7	.4	.73–.83
Financial Resources	2	3.30	1.07	.87	1.3	13.9	4.6	.88–.92
Peers and Social Support	4	3.75	.84	.88	1.7	13.3	.5	.78–.90
School Environment	4	3.62	.81	.88	.2	9.1	.4	.78–.92

^a The KIDSCREEN-52 Physical Well-being subscale and KIDSCREEN-27 Physical Well-being subscale consist of the same items

students were investigated. Participants were recruited from nine primary and 10 secondary schools in Hong Kong. Written informed consent was sought from participants and their parents prior to the start of study. Data retrieved from the first and second cohorts were used to establish the factor structure and convergent/divergent validity of KIDSCREEN scores, respectively. Procedures of the study were approved by the local institutional ethics committee. Permission to translate the KIDSCREEN was obtained from the KIDSCREEN group.

Measures

English KIDSCREEN-52 items were translated into Chinese (Cantonese) using the protocol outlined by the KIDSCREEN group [7]. To ensure that translations were age-appropriate, the translated questionnaire was pilot-tested by 12 boys and girls (8–14 years). Children were asked to interpret the instructions and items. During this process, nine items were found to contain words that were not understood by younger respondents. These items were consequently reworded using terms with very similar meanings but involve simpler language that could be understood by younger participants (please contact authors for details). The final translated KIDSCREEN-52 was administered to participants in both cohorts. KIDSCREEN-27 items were extracted from the long form of the questionnaire.

Chinese version of the four-item Family Affluence Scale (FAS) [8, 9] was used as a measure of socioeconomic status. Students with higher socioeconomic backgrounds would likely have higher quality of life; therefore, FAS scores were hypothesized to be positively related to KIDSCREEN subscale scores. The FAS was only administered to the second cohort.

The Chinese Strengths and Difficulties Questionnaire (SDQ) [10, 11] was also administered to participants in the second cohort only. The 25-item questionnaire measures children's and adolescents' difficulties (e.g., emotional symptoms) and pro-social behaviors. KIDSCREEN scores were hypothesized to be negatively associated with SDQ difficulties scores and positively related to pro-social behaviors.

Data analyses

Negatively worded KIDSCREEN items were reverse-scored, i.e., higher scores represented better well-being. Data from Cohort 1 were used to examine possible floor/ceiling effects, internal consistency, and factorial validity; while data from Cohort 2 were utilized to inspect test–retest reliability, and convergent and divergent validity. A floor (ceiling) effect was deemed present if more than 15 % of the respondents achieved the lowest (highest) possible score [12]. Subscales with Cronbach's alphas >.7

had acceptable internal consistency [13]. Confirmatory factor analysis (CFA) using a weighted least square estimator was conducted using Mplus 7.2 (Los Angeles, CA: Muthén & Muthén) [14] to examine the factorial validity of KIDSCREEN scores. Apart from Chi-square statistics, the comparative fit index (CFI), Tucker–Lewis index (TLI), and root-mean-square error of approximation (RMSEA) were used to assess model fit. Acceptable model fit should be represented by a nonsignificant Chi-square, with CFI and TLI $\geq .90$, and RMSEA $\leq .08$ [15]. A partial credit Rasch model [16] was also used to evaluate KIDSCREEN-27 responses. Analyses were conducted using Winsteps 3.74 (Beaverton, Oregon: Winsteps.com) [17] and were applied to individual subscales. Items were deemed as well fitting with infit values between .7 and 1.3. Similar to elsewhere [6], only infit statistics were presented in this study, as outfit statistics suggested similar results in terms of fit/misfit. Dimensionality maps were produced using Winsteps to determine the unidimensionality of subscale items. Test-retest reliability was assessed using intraclass correlation coefficients ICC(3,1) [18].

Results

Cohort 1—descriptive statistics, internal consistency, factorial validity

Descriptive statistics for the KIDSCREEN-52 subscale scores are shown in Table 1. KIDSCREEN-52 subscales showed acceptable Cronbach's alpha values ranging from .74 to .95. No ceiling or floor effects were observed, with the exception of the Bullying subscale where 30.1 % of participants reported the maximum score (i.e., least bullying behaviors). The hypothesized 10-factor model of the KIDSCREEN-52 was examined using CFA. Despite a significant Chi-square value, fit indices supported the 10-factor structure of KIDSCREEN-52 scores: $\chi^2(1,229) = 9389.21$, $p < .01$, CFI = .91, TLI = .90, RMSEA = .07.

KIDSCREEN-27 subscales also had acceptable internal consistency (.84–.89). The five-factor model was assessed using CFA, but the fit was poor: $\chi^2(314) = 4553.16$, $p < .01$, CFI = .91, TLI = .90, RMSEA = .10. A partial credit Rasch model was employed for each subscale to identify potentially problematic items or structures. Three items from the Psychological Well-being subscale (Items 9, 14, 17; infit = .60, 1.37, 1.37, respectively) and one from the School Environment subscale (Item 47, infit = 1.34) had poor infit statistics; all other items fitted well. Dimensionality maps (Fig. 1) suggested that the unidimensionality assumption was not met for the Psychological Well-being subscale and the Autonomy and Parent Relation subscale. Closer inspection revealed that items

representing the different dimensions were from different subscales in the KIDSCREEN-52 questionnaire. This suggested that they should be kept separate in the Chinese KIDSCREEN-27.

Analyses were repeated after separating the items as suggested by the dimensionality maps (see Fig. 2; Table 2). Specifically, the Psychological Well-being subscale was separated into the (modified) Psychological Well-being (four items) and Moods and Emotions subscales (three items). The Moods and Emotions items had good infit values, but misfitting was still evident in one item from the modified Psychological Well-being subscale (Item 9, infit = .58). This item was retained in the ongoing analyses to determine whether the misfit was sample specific. For the Autonomy and Parent Relation subscale, items were separated into the (modified) Autonomy and Parent Relation subscale (five items) and the Financial Resources subscale (two items). Items for these modified subscales had good fit. The structure of the modified Chinese KIDSCREEN-27 with seven factors was then re-examined using CFA. An acceptable model fit was found: $\chi^2(303) = 2507.43$, $p < .01$, CFI = .95, TLI = .95, RMSEA = .07.

Cohort 2—factorial validity, convergent/divergent validity, test-retest reliability

For KIDSCREEN-27 scores from the second cohort (both test and retest) using the original five-subscale structure, similar ill-fitting issues were observed from CFA and Rasch analyses. In contrast, well-fitting CFA results were found using the newly proposed seven-factor model. Similar to the results from the previous cohort, results from Rasch analyses also suggested that all but one item (Item 9) had good fit.

In terms of convergent/divergent validity, KIDSCREEN-52 and KIDSCREEN-27 subscale scores were associated with participants' socioeconomic status, psychological difficulties, and pro-social behaviors in the hypothesized direction (see Table 3). The only exception was a nonsignificant correlation between socioeconomic status and the KIDSCREEN-52 Bullying subscale scores. All KIDSCREEN-52 and KIDSCREEN-27 subscales had ICCs above .70; therefore, test-retest reliability was deemed adequate.

Discussion

In the current study, various aspects of reliability and validity of two versions of the Chinese KIDSCREEN questionnaire were examined. For the Chinese KIDSCREEN-52, the 10-factor model was supported. Evidence for convergent/divergent validity and test-retest reliability of subscale scores was also found. However, for the

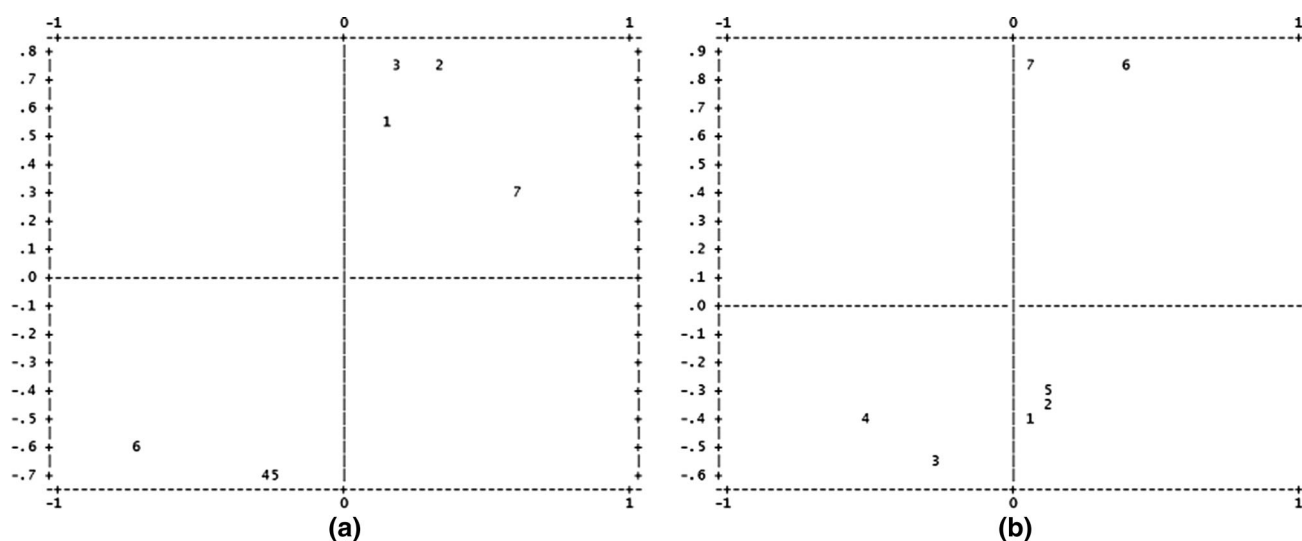


Fig. 1 Dimensionality maps for the initial KIDSCREEN-27. **a** Psychological Well-being and **b** Autonomy and Parent Relation subscale. The isolated clustering of item scores suggests that the unidimensionality assumption may not be met

Fig. 2 Subscales and corresponding items of the original and proposed chinese KIDSCREEN-27

Original KIDSCREEN Subscales	KIDSCREEN-27 Items	Proposed Chinese KIDSCREEN-27 Subscales
Physical Well-being	Item 1 Item 2 Item 3 Item 4 Item 5	Physical Well-being
Psychological Well-being	Item 6 Item 7 Item 8 Item 12	Psychological Well-being
	Item 9 Item 10 Item 11	Moods & Emotions
Autonomy & Parent Relation	Item 13 Item 14 Item 15 Item 16 Item 17	Autonomy & Parent Relation
	Item 18 Item 19	Financial Resources
Peers & Social Support	Item 20 Item 21 Item 22 Item 23	Peers & Social Support
School Environment	Item 24 Item 25 Item 26 Item 27	School Environment

KIDSCREEN-27, results suggested that seven subscales (as opposed to five) should be used. Results were consistent across all administrations (Cohort 1, test and retest of

Cohort 2) of the questionnaire. Based on the newly proposed structure of the questionnaire, support for reliability and validity of KIDSCREEN-27 scores was found.

Table 2 Cronbach's alpha, infit statistics, and factor loadings for the Chinese KIDSCREEN-27 scores in Cohorts 1 and 2 (test and retest components), using the modified seven-factor model

Item	Cohort 1 <i>N</i> = 1379 Infit ^a /factor loading ^b	Cohort 2—test <i>N</i> = 555 Infit ^a /factor loading ^b	Cohort 2—retest <i>N</i> = 526 Infit ^a /factor loading ^b
Physical Well-being	$\alpha = .84$	$\alpha = .81$	$\alpha = .85$
Item 1	.96/.68	.93/.73	.98/.72
Item 2	.70/.85	.73/.81	.75/.84
Item 3	1.21/.63	1.29/.54	1.31/.63
Item 4	1.09/.68	1.05/.64	1.00/.75
Item 5	1.01/.91	.98/.88	.94/.94
Psychological Well-being	$\alpha = .91$	$\alpha = .85$	$\alpha = .89$
Item 6	.86/.89	.82/.84	.94/.86
Item 9	.58/.92	.63/.86	.62/.93
Item 11	.86/.90	1.26/.79	1.18/.79
Item 19	1.24/.82	1.27/.76	1.23/.85
Moods and Emotions	$\alpha = .84$	$\alpha = .78$	$\alpha = .82$
Item 13	.78/.86	.85/.76	.72/.89
Item 14	1.01/.81	1.01/.77	.98/.79
Item 17	1.20/.88	1.21/.84	1.32/.80
Autonomy and Parent Relation	$\alpha = .87$	$\alpha = .82$	$\alpha = .86$
Item 24	.97/.81	1.08/.65	1.06/.77
Item 25	1.09/.76	1.01/.65	.87/.82
Item 32	.80/.83	1.03/.80	1.03/.79
Item 33	1.14/.73	.95/.76	1.15/.72
Item 34	.97/.78	.90/.76	.86/.78
Financial Resources	$\alpha = .87$	$\alpha = .80$	$\alpha = .86$
Item 35	.97/.88	.95/.83	.96/.91
Item 36	.99/.92	1.00/.89	.98/.89
Peers and Social Support	$\alpha = .88$	$\alpha = .84$	$\alpha = .87$
Item 38	1.24/.82	1.29/.78	1.20/.84
Item 40	.72/.90	.77/.85	.69/.91
Item 41	.71/.90	.66/.90	.73/.87
Item 43	1.29/.78	1.25/.71	1.35/.74
School Environment	$\alpha = .88$	$\alpha = .80$	$\alpha = .84$
Item 44	.78/.91	1.08/.88	1.08/.87
Item 45	.72/.92	1.04/.63	1.09/.71
Item 47	1.34/.78	.91/.72	.77/.81
Item 49	1.12/.78	.95/.75	1.03/.80

Item numbers corresponding to the KIDSCREEN-52 are shown; item contents are not shown due to copyright

^a Infit statistics are derived from Rasch analyses

^b Factor loadings of item scores from a seven-factor confirmatory factor analysis

Based on the seven-factor structure, items from the original KIDSCREEN-27 Psychological Well-being subscale were separated into two distinct factors. Specifically, items were separated into positively and negatively worded ones. Previous studies [19, 20] have shown that respondents, especially younger children, may be unable to accurately report their perceptions using negatively worded items. This might be the reason why items were deemed to fall into distinct factors. Regarding the original Autonomy

and Parent Relation subscale, results from Rasch analyses suggested that two items tapping the availability of financial resources were measuring a distinct construct. Although parents were likely to be students' sources of financial support, it appeared that students considered this to be independent of their relationships with parents. Nonetheless, similar issues for these subscales were not reported in previous studies [6, 21]. There is currently insufficient information to evaluate whether such differences

Table 3 Test of convergent/divergent validity and test–retest reliability of KIDSCREEN scores

	FAS	SDQ—difficulties	SDQ—pro-social behaviors	Test–retest ICC
KIDSCREEN-52 subscales				
Physical Well-being	.22**	−.35**	.35**	.86
Psychological Well-being	.18**	−.44**	.39**	.81
Moods and Emotions	.14**	−.54**	.21**	.82
Self Perception	.15**	−.46**	.22**	.80
Autonomy	.22**	−.29**	.24**	.79
Parent Relation and Home Life	.25**	−.34**	.33**	.84
Financial Resources	.27**	−.24**	.18**	.83
Peers and Social Support	.21**	−.41**	.47**	.81
School Environment	.16**	−.41**	.50**	.86
Bullying	.05	−.41**	.13**	.72
Modified KIDSCREEN-27 subscales				
Physical Well-being	.22**	−.35**	.35**	.86
Psychological Well-being	.19**	−.45**	.39**	.81
Moods and Emotions	.11*	−.50**	.17**	.78
Autonomy and Parent Relation	.24**	−.34**	.28**	.82
Financial Resources	.25**	−.22**	.15**	.80
Peers and Social Support	.20**	−.40**	.45**	.78
School Environment	.19**	−.42**	.49**	.84

FAS Family Affluence Scale, SDQ Strengths and Difficulties Questionnaire, ICC intraclass correlation

* $p < .05$; ** $p < .01$

were due to language issues, differences in cultural adaptations, or other possible factors. This should be examined in future studies.

In conclusion, evidence from the current study suggested that the Chinese version of the KIDSCREEN questionnaire is a viable measure of HRQoL in Hong Kong Chinese children and adolescents. Results based on two independent cohorts were found to be consistent, but differed slightly to the factor structure of the KIDSCREEN-27 found in previous studies using other language versions of the questionnaire.

Acknowledgments The authors would like to thank the Coca-Cola Foundation for sponsoring this study through support of a larger project.

References

1. Strine, T. W., Chapman, D. P., Balluz, L., & Mokdad, A. H. (2008). Health-related quality of life and health behaviors by social and emotional support. *Social Psychiatry and Psychiatric Epidemiology*, 43(2), 151–159.
2. The KIDSCREEN Group Europe. (2006). *The KIDSCREEN Questionnaires: Quality of life questionnaires for children and adolescents; handbook*. Lengerich: Pabst Science Publishers.
3. Ravens-Sieberer, U., Gosch, A., Rajmil, L., Erhart, M., Bruil, J., Duer, W., et al. (2005). KIDSCREEN-52 quality-of-life measure for children and adolescents. *Expert Review of Pharmacoeconomics and Outcomes Research*, 5(3), 353–364.
4. Haraldstad, K., Christophersen, K.-A., Eide, H., Nativg, G. K., & Helseth, S. (2011). Health related quality of life in children and adolescents: Reliability and validity of the Norwegian version of KIDSCREEN-52 questionnaire, a cross sectional study. *International Journal of Nursing Studies*, 48(5), 573–581.
5. Ravens-Sieberer, U., Gosch, A., Rajmil, L., Erhart, M., Bruil, J., Power, M., et al. (2008). The KIDSCREEN-52 quality of life measure for children and adolescents: Psychometric results from a cross-cultural survey in 13 European countries. *Value in Health*, 11(4), 645–658.
6. Robitail, S., Ravens-Sieberer, U., Simeoni, M. C., Rajmil, L., Bruil, J., Power, M., et al. (2007). Testing the structural and cross-cultural validity of the KIDSCREEN-27 quality of life questionnaire. *Quality of Life Research*, 16(8), 1335–1345.
7. The KIDSCREEN Group. (2004). Translation & validation procedure: Guidelines and documentation form.
8. Wardle, J., Robb, K., & Johnson, F. (2002). Assessing socioeconomic status in adolescents: The validity of a Home Affluence Scale. *Journal of Epidemiology and Community Health*, 56(8), 595–599.
9. Liu, Y., Wang, M., Villberg, J., Torsheim, T., Tynjälä, J., Lv, Y., & Kannas, L. (2012). Reliability and validity of Family Affluence Scale (FAS II) among adolescents in Beijing, China. *Child Indicators Research*, 5(2), 235–251.
10. Goodman, R., Meltzer, H., & Bailey, V. (1998). The Strengths and Difficulties Questionnaire: A pilot study on the validity of the self-report version. *European Child and Adolescent Psychiatry*, 7(3), 125–130.
11. Lai, K. Y. C., Luk, E. S. L., Leung, P. W. L., Wong, A. S. Y., Law, L., & Ho, K. (2010). Validation of the Chinese version of the strengths and difficulties questionnaire in Hong Kong. *Social Psychiatry and Psychiatric Epidemiology*, 45(12), 1179–1186.
12. Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., et al. (2007). Quality criteria

- were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60(1), 34–42.
13. Reeve, B. B., Wyrwich, K. W., Wu, A. W., Velikova, G., Terwee, C. B., Snyder, C. F., et al. (2013). ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Quality of Life Research*, 22(8), 1889–1905.
 14. Muthén, L. K., & Muthén, B. O. (2008–2011). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthén and Muthén.
 15. Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(2), 23–74.
 16. Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). New York: Routledge.
 17. Linacre, J. M. (2014). *Winsteps®* (Version 3.74) [Computer software]. Beaverton, Oregon: Winsteps.com.
 18. Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine*, 30(1), 1–15.
 19. DiStefano, C., & Motl, R. W. (2006). Further investigating method effects associated with negatively worded items on self-report surveys. *Structural Equation Modeling*, 13(3), 440–464.
 20. Barnette, J. J. (2000). Effects of stem and likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement*, 60(3), 361–370.
 21. Ravens-Sieberer, U., Herdman, M., Devine, J., Otto, C., Bullinger, M., Rose, M., & Klasen, F. (2014). The European KIDSCREEN approach to measure quality of life and well-being in children: Development, current application, and future advances. *Quality of Life Research*, 23(3), 791–803.