Is the Kidscreen-27 a valid measure of health-related quality of life in 10-year-old Norwegian children?

John Roger Andersen, Gerd Karin Natvig, Kristin Haraldstad, Turid Skrede, Eivind Aadland, Geir Kåre Resaland

The aim of this study was to investigate the reliability and validity of the Norwegian Kidscreen-27 questionnaire, a measure of generic health-related quality of life, in 10 year-old children. The Kidscreen-27 consists of five domains and was validated in a sample of 56 school children (29 boys). The children completed the questionnaire at three different time points during two consecutive school days. For convergent validity, the study was powered to detect a statistically significant correlation coefficient of 0.4. Cronbach's alpha values ranged from 0.73 to 0.83. Floor effects were all zero and ceiling effects ranged from 1.7% to 23.7%. Intraclass correlation values over time ranged from 0.71 to 0.81. However, some individual variability over time occurred and was illustrated by Bland Altman plots. The domains of physical well-being, psychological well-being and autonomy & parents improved over time (Ps < 0.05), while social support and school environment domains did not. We assessed convergent validity using general life satisfaction scores obtained by administering the Cantrils Ladder. All the Kidscreen-27 domains were significantly associated with general life satisfaction (Spearman rank correlations ranged from 0.29 to 0.59, Ps < 0.05). In conclusion, the Norwegian version of Kidscreen-27 has good reliability and validity.
Is the Kidscreen-27 a valid measure of health-related quality of life in 10-year-old Norwegian children?

John Roger Andersen 1,2, Gerd Karin Natvig 3, Kristin Haraldstad 4, Turid Skrede 5, Eivind Aadland 5, Geir Kåre Resaland 2,5

1 Department of health studies, Sogn og Fjordane University College, Førde, Norway
2 Centre of Health Research, Førde Hospital Trust, Førde, Norway
3 Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway
4 Department of Health and Nursing Sciences, University of Agder, Kristiansand, Norway
5 Faculty of Teacher Education and Sport, Sogn og Fjordane University College, Sogndal, Norway

Correspondence. John Roger Andersen. Department of Health Studies, Sogn og Fjordane University College, Førde, Norway. Address: Svanehaugvegen 1, 6812 Førde, Norway. Telephone: +47 57722522. E-mail: johnra@hisf.no
Abstract

The aim of this study was to investigate the reliability and validity of the Norwegian Kidscreen-27 questionnaire, a measure of generic health-related quality of life, in 10 year-old children. The Kidscreen-27 consists of five domains and was validated in a sample of 56 school children (29 boys). The children completed the questionnaire at three different time points during two consecutive school days. For convergent validity, the study was powered to detect a statistically significant correlation coefficient of 0.4. Cronbach's alpha values ranged from 0.73 to 0.83. Floor effects were all zero and ceiling effects ranged from 1.7% to 23.7%. Intraclass correlation values over time ranged from 0.71 to 0.81. However, some individual variability over time occurred and was illustrated by Bland Altman plots. The domains of physical well-being, psychological well-being and autonomy & parents improved over time (Ps < 0.05), while social support and school environment domains did not. We assessed convergent validity using general life satisfaction scores obtained by administering the Cantrils Ladder. All the Kidscreen-27 domains were significantly associated with general life satisfaction (Spearman rank correlations ranged from 0.29 to 0.59, Ps < 0.05). In conclusion, the Norwegian version of Kidscreen-27 has good reliability and validity.

Background

The Kidscreen-27 is a well-validated, short, multidimensional measure of generic health-related quality of life (HRQoL) in children and adolescents (Ravens-Sieberer 2006; Ravens-Sieberer et al. 2007). It is available in many languages, including Norwegian. However, unlike the longer Kidscreen-52, the Norwegian version of the Kidscreen-27 has not yet been validated (Haraldstad et al. 2011). The Kidscreen-27 is a particularly useful instrument for assessing HRQoL in younger children as a consequence of its content, ease of completion and the volume of available data with which international comparisons can be made. The aim of this study was to investigate the reliability and validity of the Norwegian Kidscreen-27 questionnaire in 10-year-old children.

Material and Methods

This was a prospective validation study, and was approved by the regional ethical research committee (2012/1089). We obtained written informed consent from the parents or guardians of the participants. We used G-POWER to calculate that a minimum sample size of 44 children would be required. This assumed an alpha value of 0.05, a power of 0.80, and a minimum correlation of 0.4 for a two-tailed test. The proposed effect size was based on reported correlations between the Kidscreen-27 and other quality of life measures (Ravens-Sieberer 2006; Ravens-Sieberer et al. 2007). We recruited children aged 10 years from the same school in Western Norway.
The Kidscreen-27 was administered to the children at three time points: at the beginning of the school day (Test 1); at the end of the same school day (Test 2); and during the middle of the following school day (Test 3). The children were given brief information about the questionnaire, and then completed it while they sat at their desks in the classroom. After doing so, they were told to carry on with their school work in order to minimize any noise until all the other children were finished. They were allowed to ask the teacher for help if they did not understand any of the questions. Children with reading difficulties were helped by a teaching assistant, as would be the case normally.

**Kidscreen-27**

Items of the Kidscreen-27 are derived from the Kidscreen-52 questionnaire (Ravens-Sieberer 2006). It has five domains: (1) physical well-being; (2) psychological well-being; (3) autonomy & parents; (4) social support & peers; and (5) school environment. We used the methodology given in the manual to obtain the T-scores; mean (±SD) scores of 50 ± 10 define normality for children and adolescents aged 8-18 years across Europe (Ravens-Sieberer 2006). Higher scores indicate a better HRQoL. The Kidscreen-27 is standardized so that a difference of <0.2 points is considered trivial, 2.0 – 4.9 as small, 5.0 - 7.9 as moderate and ≥8 as large effects (Cohen 1988; Ravens-Sieberer 2006). There are two official forms of written Norwegian, Bokmål and Nynorsk. They are very similar, and it is mandatory that they are both taught in Norwegian schools. The Kidscreen-52 questionnaire is only available in Bokmål (Haraldstad et al. 2011). We decided to use Nynorsk in this study. The linguistic skills required to produce a Nynorsk version of
Kidscreen-27 from the Kidscreen-52 are minimal. The translation was undertaken by the first three authors and modified after discussion with a professor of Norwegian languages (see acknowledgements). This process was approved by the European Kidscreen Group.

Cantrils Life Satisfaction Ladder

In order to study convergent validity we assessed general life satisfaction using the adapted version of the Cantrils Life Satisfaction Ladder. This has been used in World Health Organisation surveys of children and adolescents, including in Norway (Levin & Currie 2014). The child is presented with a picture of a ladder with steps ranging from 0 to 10. They are told that the top step (10) represents the best possible quality of life, while the bottom (0) represents the worst. They are asked to indicate where on the ladder they currently consider themselves to be. This question was assessed three times together with the Kidscreen-27.

Statistical analysis

Continuous variables were given as means ± standard deviations (SDs). Categorical variables were given as numbers and percentages. An independent t-test was used to study gender differences in the Kidcreen-27 scores (Test 1). Internal consistency was assessed by calculating Cronbach’s alpha values (Test 1); values ≥0.7 were considered satisfactory (Ravens-Sieberer 2006; Ravens-Sieberer et al. 2007). Floor and ceiling effects were demonstrated by the numbers and percentages of children with the lowest and highest possible scores (Test 1). Test-retest
reliability (Tests 1 - 3) was assessed by calculating single measures intraclass correlation 
coefficients (ICC), using a two-way mixed model with an absolute agreement definition; ICC 
values ≥0.7 were considered satisfactory (Ravens-Sieberer 2006; Ravens-Sieberer et al. 2007).

Individual variability was described by presenting mean differences ± SDs and 95% limits of 
agreement; Bland Altman plots were used to graphically display the variation (Tests 1 - 3). A 
linear mixed model based on restricted maximum likelihood estimation with random intercept 
for subjects was used for analysis of change in Kidscreen-27 scores over the three time points 
(Tests 1 - 3).

For convergent validity, we used the Spearman rank correlation (r_s) to test whether the 
Kidscreen-27 scores were positively correlated with the Cantrils Ladder score (test 1). Based on 
previous research, we hypothesized that the Kidscreen-27 psychological well-being domain 
would be the one most strongly correlated with the Cantrils Ladder score (Ravens-Sieberer 
2006; Ravens-Sieberer et al. 2007). Correlation coefficients <0.1 were considered trivial, 0.1 – 
0.29 as small, 0.30 – 0.49 as moderate, and ≥0.5 as high(Cohen 1988). The software Prism 
version 6.05 for Windows was used to calculate and display the results shown in Table 3 and 
Figure 1. Other statistical analyses were performed using IBM SPSS version 21 for Windows.

Two-sided P-values <0.05 were considered statistically significant.

Results

A total of 56 children (29 boys) were present at school when the study was conducted. This 
represents an attendance rate of 88%; all participated in the study. Boys scored significantly
higher than girls on physical well-being, autonomy & parents, and social support & peers (Table 1). Table 2 shows the Kidscreen-27 mean T-value scores, the percentages of children who had floor and ceiling scores, and Cronbach’s alpha values. Reliability for the various domain scores ranged from 0.73 to 0.81. Floor effects were all zero and ceiling effects ranged from 1.7% to 23.7%. The variability over time in the five Kidscreen-27 domain scores is shown in Table 3. By way of an illustrative example, Figure 1 shows Bland Altman plots for the physical well-being domain. Bland Altman plots for the other Kidscreen-27 domains can be produced from the dataset which we have made publically available. Trends in the domain scores over time between the three tests are shown in Table 4. Physical well-being, psychological well-being and autonomy & parents significantly improved over time.

The Kidscreen-27 scores were positively and significantly correlated with the Cantrils Ladder scores. The correlations for the various Kidscreen-27 domains were as follows (N = 55 to 56):

- physical well-being, $r_s = 0.29\ (P = 0.031)$;
- psychological well-being, $r_s = 0.59\ (P < 0.001)$;
- autonomy & parents, $r_s = 0.31\ (P = 0.025)$;
- social support & peers, $r_s = 0.53\ (P < 0.001)$;
- school environment, $r_s = 0.48\ (P < 0.001)$.

Discussion

The findings of this study demonstrate that the reliability and validity of the Norwegian version of the Kidscreen-27 are good. A large body of research findings regarding the Kidscreen-27 enables a direct comparison with our current results (Ravens-Sieberer 2006; Ravens-Sieberer et al. 2007). We found that boys had higher scores than girls on three of the five Kidscreen-27
domains. This is in accordance with previous findings (Ravens-Sieberer 2006; Ravens-Sieberer et al. 2007). Furthermore, the Cronbach alpha values in the current study ranged from 0.73 to 0.83. This compares to 0.80 to 0.84 reported in the literature when the instrument was administered to children aged between 8 and 18 years (Ravens-Sieberer 2006). We found no floor effects, but a moderate ceiling effect in the social support & peers domain. This is similar to previously reported findings (Ravens-Sieberer 2006).

The ICC values in the current study ranged from 0.72 to 0.80, which were slightly higher than the range of 0.61 to 0.66 reported previously (Ravens-Sieberer et al. 2007). This finding may be a consequence of different test-retest intervals (1 day vs. 2 weeks). Additionally, our study showed variability in Kidscreen-27 scores over time in some individuals. This has been reported with other HRQoL measures and in various populations (van der Velde et al. 2009; Wink et al. 2013). Although it could be a consequence of measurement error, natural fluctuations in HRQoL could also be responsible (Fayers & Machin 2007). Previous research has suggested that there is a small retest effect, in order that Kidscreen-27 scores tend to rise if assessed multiple times, even without any interventions (Ravens-Sieberer 2006). Our study demonstrated this in three of the domains, even though the test-retest intervals were short. Whether this effect is still present if the test-retest interval is longer, is unknown. It means that studies using Kidscreen-27 to assess the effectiveness of interventions must have a control group.

Finally, we found that all the Kidscreen-27 domains were significantly associated with general life satisfaction as measured with the Cantrils Ladder. The sizes of the correlation coefficients are similar to those that have been reported previously when assessing convergent validity.
using a range of instruments and in various populations (Ravens-Sieberer 2006; Ravens-
Sieberer et al. 2007).

This study was limited as only children aged 10 years were included. It is more usual to assess a
wider age group. Our decision to do so was a pragmatic one, as we required a validated
instrument to assess HRQoL in 10-year-olds, for the purpose of a randomized controlled trial.

However, we propose that if good reliability and validity levels of an assessment instrument are
demonstrated in younger children, it is likely that these will also be the case in older children
and adolescents. Our short test–retest intervals might be criticized. However, in addition to the
points made above, it should be noted that a previous study found that test-retest effects were
similar whether the interval is 2 days or 2 weeks (Marx et al. 2003).

Conclusions
Our findings suggest that the Kidscreen-27 works well in Norwegian context, and has good
reliability and validity. Further larger studies are needed to assess the Kidscreen-27 more fully
with regards to its clinical and research utility, and its ability to detect changes in HRQL
following interventions.

Acknowledgements
We thank Associate Professor Jan Olav Fretland from the Faculty of Teacher Education and
Sport, Sogn og Fjordane University College, for his assistance with the preparation of the
‘Nynorsk’ Kidscreen-27.
References


Table 1: Physiological measurements, Kidscreen-27 domain scores, and Cantrils Ladder scores of the participants

These results are from Test 1. Variables are means ± SD.
<table>
<thead>
<tr>
<th>Variables</th>
<th>Boys (n = 29)</th>
<th>Girls (n = 27)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical well-being (n = 55)</td>
<td>51.4 ± 9.8</td>
<td>46.6 ± 6.7</td>
<td>0.039</td>
</tr>
<tr>
<td>Psychological well-being (n = 56)</td>
<td>53.3 ± 8.7</td>
<td>51.0 ± 7.9</td>
<td>0.305</td>
</tr>
<tr>
<td>Autonomy &amp; parents (n = 55)</td>
<td>52.6 ± 8.6</td>
<td>47.5 ± 6.0</td>
<td>0.014</td>
</tr>
<tr>
<td>Social support &amp; peers (n = 56)</td>
<td>56.2 ± 9.5</td>
<td>50.9 ± 9.4</td>
<td>0.043</td>
</tr>
<tr>
<td>School environment (n = 56)</td>
<td>55.3 ± 10.8</td>
<td>52.6 ± 8.7</td>
<td>0.301</td>
</tr>
<tr>
<td>Cantrils Ladder score (n = 56)</td>
<td>8.9 ± 1.5</td>
<td>8.1 ± 2.1</td>
<td>0.152</td>
</tr>
</tbody>
</table>
Table 2: Kidscreen-27 T-value domain scores, floor and ceiling effects, and internal consistency

These results are from Test 1. Variables are means ± SD.<!-��;WqW-->
<table>
<thead>
<tr>
<th>Domains</th>
<th>T-value</th>
<th>% Floor</th>
<th>% Ceiling</th>
<th>Cronbach’s alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical well-being (N= 55)</td>
<td>49.1 ± 9.0</td>
<td>0</td>
<td>3.4</td>
<td>0.73</td>
</tr>
<tr>
<td>Psychological well-being (N = 56)</td>
<td>52.3 ± 8.7</td>
<td>0</td>
<td>3.4</td>
<td>0.73</td>
</tr>
<tr>
<td>Autonomy &amp; parents (N = 55)</td>
<td>50.0 ± 7.8</td>
<td>0</td>
<td>1.7</td>
<td>0.79</td>
</tr>
<tr>
<td>Social support &amp; peers (N = 56)</td>
<td>53.5 ± 9.5</td>
<td>0</td>
<td>23.7</td>
<td>0.83</td>
</tr>
<tr>
<td>School environment (N = 56)</td>
<td>54.2 ± 9.9</td>
<td>0</td>
<td>13.6</td>
<td>0.81</td>
</tr>
</tbody>
</table>
Table 3: Variation in the Kidscreen-27 scores over time

The results are from Tests 1 - 3.
<table>
<thead>
<tr>
<th>Domains</th>
<th>Differences ± SD</th>
<th>95% limits of agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Physical well-being</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test 2 - 1 (N = 52)</td>
<td>1.9 ± 7.3</td>
<td>-12.4, 16.2</td>
</tr>
<tr>
<td>Test 3 - 1 (N = 51)</td>
<td>5.4 ± 8.6</td>
<td>-11.47, 22.3</td>
</tr>
<tr>
<td>Test 3 - 2 (N = 49)</td>
<td>2.8 ± 6.0</td>
<td>-8.9, 14.5</td>
</tr>
<tr>
<td><strong>Psychological well-being</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test 2 - 1 (N = 56)</td>
<td>2.5 ± 7.1</td>
<td>-11.4, 16.3</td>
</tr>
<tr>
<td>Test 3 - 1 (N = 51)</td>
<td>2.2 ± 8.5</td>
<td>-14.5, 18.7</td>
</tr>
<tr>
<td>Test 3 - 2 (N = 51)</td>
<td>-0.1 ± 6.9</td>
<td>-13.6, 13.4</td>
</tr>
<tr>
<td><strong>Autonomy &amp; parents</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test 2 - 1 (N = 54)</td>
<td>1.2 ± 5.3</td>
<td>-9.1, 11.5</td>
</tr>
<tr>
<td>Test 3 - 1 (N = 50)</td>
<td>3.8 ± 8.3</td>
<td>-12.4, 19.9</td>
</tr>
<tr>
<td>Test 3 - 2 (N = 50)</td>
<td>2.5 ± 7.7</td>
<td>-12.7, 17.6</td>
</tr>
<tr>
<td><strong>Social support &amp; peers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test 2 - 1 (N = 56)</td>
<td>-0.9 ± 4.9</td>
<td>-10.5, 8.7</td>
</tr>
<tr>
<td>Test 3 - 1 (N = 53)</td>
<td>-1.4 ± 5.9</td>
<td>-13.0, 10.2</td>
</tr>
<tr>
<td>Test 3 - 2 (N = 53)</td>
<td>-0.6 ± 8.5</td>
<td>-12.1, 10.9</td>
</tr>
<tr>
<td><strong>School environment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test 2 - 1 (N = 55)</td>
<td>0.6 ± 4.4</td>
<td>-8.1, 9.3</td>
</tr>
<tr>
<td>Test 3 - 1 (N = 50)</td>
<td>1.0 ± 8.4</td>
<td>-15.4, 17.4</td>
</tr>
<tr>
<td>Test 3 - 2 (N = 49)</td>
<td>0.4 ± 6.8</td>
<td>-12.9, 13.8</td>
</tr>
</tbody>
</table>
Table 4: Time trends for changes in the Kidscreen-27 domain scores

The results are from Tests 1 - 3. The scores are presented as mean change values and 95% CIs.
<table>
<thead>
<tr>
<th>Domains</th>
<th>Change (Test 2 - 1)</th>
<th>Change (Test 3 - 1)</th>
<th>P for trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical well-being (N= 55)</td>
<td>2.1 (0.1, 4.1)</td>
<td>5.3 (3.3, 7.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Psychological well-being (N = 56)</td>
<td>2.5 (0.5, 4.5)</td>
<td>2.3 (0.3, 4.4)</td>
<td>0.025</td>
</tr>
<tr>
<td>Autonomy &amp; parents (N = 55)</td>
<td>1.1 (-0.8, 3.1)</td>
<td>3.8 (1.8, 5.8)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Social support &amp; peers (N = 56)</td>
<td>-0.9 (-2.3, 0.6)</td>
<td>-1.4 (-2.9, 0.1)</td>
<td>0.188</td>
</tr>
<tr>
<td>School environment (N = 56)</td>
<td>0.6 (-1.1, 2.4)</td>
<td>1.1 (-0.7, 2.9)</td>
<td>0.480</td>
</tr>
</tbody>
</table>
Figure 1: Bland Altman plots

Bland Altman plots showing differences between: Tests 1 and 2 (A); 1 and 3 (B); and 2 and 3 (C), as a function of the mean of the corresponding tests on the Kidscreen-27 physical well-being domain.
A. Average scores of test 1 and 2

Difference (test 2 - 1)

B. Average scores of test 1 and 3

Difference (test 3 - 1)

C. Average scores of test 2 and 3

Difference (test 3 - 2)