The impact of paediatric inflammatory bowel disease. Epidemiology, disease activity and quality of life

Loonen, H.J.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 3.4

A Comparison of Likert Scale and Visual Analogue Scales in Children

Henriëtte van Laerhoven, Hester J van der Zaag – Loonen, and Bert HF Derkx
Abstract

Objective
To examine which type of response options children find easiest to perform, which response option they prefer, and to study the relative validity of the different response options.

Methods
A consecutive group of 120 children aged 6-18 years old filled out a 7-item questionnaire including two affective questions, three opinion questions and two frequency questions. They completed three pages with different answering modes in random order (Likert scale, simple Visual Analogue Scale (VAS) and numeric VAS). They gave a mark for ease and preference of each response option.

Results
Median marks for ease were 10.0 for the Likert scale, 7.5 for the simple VAS and 9.0 for the numeric VAS. Median marks for preference were: 9.0 for the Likert scale, 6.0 for the simple VAS and 8.0 for the numeric VAS. Correlation coefficients as a measure of relative validity of the response options were generally high (> 0.80) between all response options.

Conclusions
Children regardless of age find the Likert scale the easiest and the most preferable answering option. Overall, the Likert scale, the simple VAS and the numeric VAS show comparable validity. Therefore, the Likert scale is recommended for use in questionnaires for children.
Introduction

Most methodological knowledge on instruments to measure for example health status and quality of life (QoL), goes out to research on adults. There is a relative lack of methodological knowledge in children.\(^1\) So far most researchers draw upon personal experiences. However, there is an increasing interest in measuring childhood QoL, both in clinical practice as well as in clinical studies. Specific methodological issues encountered in children are the minimum skills of children required to be able to complete instruments, age cut-offs versus cut-offs defined by the developmental state of the child, and preventing the occurrence of missing data.\(^2\)

Considerable work in adults has gone into examining the use of various response options. The verbal categorical response option (i.e. discrete verbal answering options also known as Likert scale response options) and analogue scales (i.e. a visual analogue scale called VAS, or a numeric scale anchored at each end with the extreme opposites) are most widely used. Both methods proved to be convenient and valid to quantify the outcome pain severity.\(^3\) The VAS is however acknowledged as the best response option to assess pain severity since it does not force quantum changes as occur with the Likert scale. For the assessment of other outcomes the Likert response scale is often used because of the few cognitive demands it places upon patients, and because of ease of interpretation for the investigator.\(^4,6,8-10\)

There are advantages and disadvantages to both types of response options. Some authors favour the VAS because it seems to be more sensitive to small changes and is thought to increase answering precision.\(^4,8\) However, this precision is often illusory, because the meaning of a one-centimetre difference is unknown and may vary from person to person. Another disadvantage is that the VAS is cognitively more demanding and may therefore require explanation and training before patients are able to use it.\(^4,8-10\) The Likert scale is recommended by others because of its ease of interpretation, and because the significance of a change of one or more points on the Likert scale seems easier to grasp intuitively.\(^8,9\) For the Likert scale neither training nor mathematical background are required.\(^3,4,6,8,10\) The disadvantage of the Likert scale is that it forces a choice of existing answers.\(^8\) Furthermore, the number of responses chosen by the instrument developers dictates the amount of measurement information obtained.

In research on adults a substantial correlation between the two different response options has been found.\(^3,4,6,9\) Most authors agree that there are few differences in reliability and responsiveness between the response options and prefer the Likert scale because of ease of completion and score computation.\(^3,6,8,9,11\)

Generally children as old as eight years are believed to be able to reliably report on their well being.\(^12\) Some authors favour the use of standardized reading tests instead of age limits to provide guidance as to the minimum skills needed to complete questionnaires.\(^13\) Little is
known about the ability of younger children to complete self report instruments. As can be appreciated, the cognitive and developmental level of children makes them different from adults. Therefore, methodology obtained from research on adults cannot simply be transposed to the paediatric population. Although few empirical data regarding how adept children of different ages are in using Likert response options, Likert scales are most often used in paediatric QoL measures. Research on elderly people and illiterate persons showed them to experience difficulties in completing instruments that use the VAS. When one regards children cognitively comparable to these populations, they may be expected to experience the same difficulties when using the VAS. Others say children may fail to use the full range of the VAS scale available, and notably focus on the endpoints, thereby decreasing the amount of measurement potential. On the other hand, children have less reading potential, and may therefore be expected to prefer the response options according to the VAS, because it requires less reading time and skills than the Likert response options. However, there is little information about the way children handle different response options.

Therefore the purpose of this study was to examine which response options children find easiest to complete and which response option they prefer. We also studied the correlation between the different response options.

**Methods**

A questionnaire containing seven items was developed, asking three types of questions: affection, opinion and frequency. Affection was measured with two questions asking about dreams and current mood. Three questions asked about the child’s opinion on school, sports and their height. The remaining two questions asked how often the child watches television and how often he goes by bus.

A consecutive group of children between 6 and 18 years old visiting a general paediatric outpatient clinic was given a three page questionnaire package. One page contained the seven questions and the verbal categorical response option (Likert scale) as answering mode. On this version the response options were verbally expressed as five possible answers. (Figure 1a) A second page with the same questions had the numeric VAS as answering option. On this page the response options were framed as a series of numbers from one to ten with at each end of the line the two extreme answers from the Likert scale. (Figure 1b) A third page had the VAS as answering mode. This version presented a 10 cm line with again the same anchors on both ends of the line. The child was asked to give the answer by putting a cross on the line between the two extreme answers. (Figure 1c) In addition to the instructions given by the investigator each page contained written instructions. The three pages were presented in random order and the children were asked to fill in the questionnaires themselves.

Prior to completing the package, the children were asked questions concerning age, sex
and the type of school they are attending. Ethnic background was registered and the time to complete the package of questionnaires was recorded.

After completing all pages, children were asked to rate the ease and preference for each answering option by giving them a mark from one to ten, comparable to marks given in school in the Netherlands.

**Statistical analysis**

Marks given for ease and preference of each response option were treated as ordinal data. Therefore, median scores were calculated and non-parametric procedures used. Differences in median ease and preference scores within patients for the three response options were compared using the Wilcoxon statistic for paired ordinal variables.

Children were categorized into three groups by age: children six and seven years old (very young children), children eight to twelve years old (young children) and those thirteen years and older (older children). Differences between age groups regarding marks for ease and preference of the response options were tested with non-parametric procedures comparing three independent groups (Kruskal Wallis test). Differences in ease and preference marks between native and immigrant children, and between boys and girls were compared using the Mann-Whitney-U test.

The correlation between the Likert and the visual analogue measurement techniques was assessed by calculating the Spearman rank order coefficient between the three response options.

Since one of the advantages of the VAS measurement technique is greater potential in score ranges, where the disadvantage in children has been suggested to be their tendency to stick to the first (left) answer, we performed additional analyses. We compared the percentage of children scoring in each category of the Likert scale with the percentage of children scoring in 5 categories of the simple VAS. We therefore categorized the total VAS score range of 0-100 into scores 0-20 into category 1, 21-40 into category 2 etc. Wilcoxon rank statistics for paired ordinal data were calculated to test differences in categories.

A conventional significance level of 0.05 was chosen.

**Results**

One hundred and twenty-two children were asked to participate, of whom two refused participation, resulting in a response rate of 99%. From these 120 children, 14 (12%) were 6 or 7 years old, 55 (46%) between 8 and 12 years old, and 51 (43%) were between 13-18 years old. Fifty-seven (48%) were boys; 44% of the children were immigrants.

The simple and the numeric VAS had more missing data than the Likert scale. They had respectively 3% and 2% missing data, while the Likert scale counted only 0.5% missings. The
misssings all occurred in the group of children older than 8 years old.

Table 1 shows the median marks for ease and preference of the different response options. There was no ordering effect of the different questionnaire pages upon the ease and preference scores. Wilcoxon’s statistic showed the Likert scale to be scored significantly higher on both ease and preference, followed by the numeric VAS and the VAS (all p values < 0.005). (Table 1) There was no difference in ease or preference score between the very young, young and older children. There was no difference between boys and girls in ease or preference scores. The immigrant children liked the simple VAS better than the native children (median score 7.0 versus 6.0, p= 0.003), and they found the simple VAS easier to fill out (9.0 versus 7.0, p= 0.001). Only 10 children in our sample attended a special school and we considered this group too small to perform separate analyses on this group.

**Table 1.**

<table>
<thead>
<tr>
<th></th>
<th>Median mark</th>
<th>Kruskal Wallis test(^c): p value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ease</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likert scale</td>
<td>10.0(^a)</td>
<td>0.10</td>
</tr>
<tr>
<td>simple VAS</td>
<td>7.5(^b)</td>
<td>0.89</td>
</tr>
<tr>
<td>numeric VAS</td>
<td>9.0</td>
<td>0.68</td>
</tr>
<tr>
<td><strong>Preference</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likert scale</td>
<td>9.0(^c)</td>
<td>0.21</td>
</tr>
<tr>
<td>simple VAS</td>
<td>6.0(^d)</td>
<td>0.81</td>
</tr>
<tr>
<td>numeric VAS</td>
<td>8.0</td>
<td>0.61</td>
</tr>
</tbody>
</table>

\(^a\) Higher than the mark for the simple VAS (p< 0.001) and numeric VAS (p= 0.005)
\(^b\) Lower than the mark for the numeric VAS (p< 0.001)
\(^c\) Higher than the mark for the simple VAS (p< 0.001) and numeric VAS (p< 0.001)
\(^d\) Lower than the mark for the numeric VAS (p< 0.001)
\(^e\) Differences in median marks for the three response options between the three age groups (children 6-7 years old, 8 to 12 years old, and 13 years and older).

The questionnaires were filled out in 2-22 minutes, with 75% of children needing between 5-10 minutes and a mean time of 7.8 minutes. On average very young children needed significantly more time to fill out the questionnaires than the young and older children: 10.9 minutes, while young children needed 7.7 minutes, and older children needed only 7.0 minutes (ANOVA test: p= 0.003). Immigrant children needed more time to fill in the questionnaires: 8.8 minutes versus 7.2 minutes (p= 0.052, 95% CI: 0.0-3.1)
Chapter 3.4 Likert versus VAS

Table 2.
Correlations between the response options for the seven questions.

<table>
<thead>
<tr>
<th></th>
<th>Dreams</th>
<th>Teachers in school</th>
<th>Frequency of watching TV</th>
<th>Current mood</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>0.67</td>
<td>0.71</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>0.73</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Sports</th>
<th>Height</th>
<th>Frequency of going by bus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>0.73</td>
<td>0.79</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>0.80</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Spearman rank order correlation coefficients are presented. A= Likert scale, B= simple VAS, C= numeric VAS.

Table 2 shows the correlation coefficients (r) between the response options for the seven questions. Overall the numeric VAS correlated with the Likert scale with a coefficient of 0.82, and with the simple VAS with a coefficient of 0.80. The simple VAS and the Likert scale correlated with a coefficient of 0.76. The hypothesis that younger children give less reliable answers would lead to expectations of lower correlation coefficients for the younger groups than for the older groups of children. Table 3 shows the coefficients for the very young children compared with the other (young and older) children. Remarkably low coefficients (<0.40) were found for one affection question (my dreams are usually), one opinion question (I am satisfied with my height), and both frequency questions (how often do you watch television, and how often do you go by bus). However, the above mentioned results are difficult to interpret because there were only fourteen six and seven year old children, therefore correlation coefficients can be highly influenced by individual outliers.

Table 4 shows the comparison of the percentage of children scoring in each category of the Likert scale and in each of the even categories of the simple VAS. Because a very high percentage of children consistently scored in the first category of the VAS, five alternative categories were formed. The two most extreme categories were made smaller (incorporating only 10 points each), and the remaining 80 points on the VAS scale were divided into the remaining three categories, resulting in an uneven categorisation of the VAS. (Table 4)
Part III Quality of Life: Methodology

Table 3.
Correlation coefficients between the response options by age.

<table>
<thead>
<tr>
<th></th>
<th>Children 6 and 7 years old (n= 14)</th>
<th>Children 8 years and older (n=106)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likert versus simple VAS</td>
<td>0.51</td>
<td>0.77</td>
</tr>
<tr>
<td>Likert versus numeric VAS</td>
<td>0.62</td>
<td>0.83</td>
</tr>
<tr>
<td>Simple VAS versus numeric VAS</td>
<td>0.54</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Average coefficients for all seven questions are presented.

Table 4.
Percentage scoring in the different categories in the Likert scale and the simple visual analogue scale (VAS).

<table>
<thead>
<tr>
<th>Item</th>
<th>% in Likert categories a</th>
<th>% in even VAS categories b</th>
<th>% in uneven VAS categories c</th>
<th>W d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dreams</td>
<td>1 35 43 3 3</td>
<td>0-20 28 21 4 5</td>
<td>0-10 23 32 5 4</td>
<td>0.001</td>
</tr>
<tr>
<td>Liking teachers</td>
<td>16 32 40 7 6</td>
<td>21-40 24 6 8</td>
<td>11-36 37-63 64-90 91-100</td>
<td>0.04</td>
</tr>
<tr>
<td>TV watching</td>
<td>31 36 18 15 0</td>
<td>41-60 5 2</td>
<td>64-90 11-36 37-63 64-90 91-100</td>
<td>0.01</td>
</tr>
<tr>
<td>Current mood</td>
<td>23 48 19 8 3</td>
<td>61-80 7 6</td>
<td>64-90 21 6 1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Playing sports</td>
<td>53 31 13 3 2</td>
<td>81-100 1 6</td>
<td>58-91 22 7 5</td>
<td>0.58</td>
</tr>
<tr>
<td>Height</td>
<td>30 46 13 8 3</td>
<td>11-36 13 2</td>
<td>11-36 22 7 5</td>
<td>0.001</td>
</tr>
<tr>
<td>Taking the bus</td>
<td>17 3 5 28 47</td>
<td>11-36 13 2</td>
<td>11-36 22 7 5</td>
<td>0.009</td>
</tr>
</tbody>
</table>

a 1-5: answering options for the Likert scale, 1 = most positive answer, 5 = most negative answer
b 0-20 etc.: even distribution of the 10-centimetre line of the simple VAS
c 0-10 etc.: uneven distribution of the 10-centimetre line of the simple VAS
d Wilcoxon statistic's p value between children scoring in categories of the Likert compared with those scoring in the evenly distributed VAS, or in the unevenly distributed VAS, respectively.

Discussion

In this study we looked at the use of different response options in questionnaires for children. Since preference is an important feature of an instrument's feasibility, we asked the children what options they preferred. The results of this study show that both the Likert scale and the visual analogue scales correlate highly and can be used as response options for children. Both younger and older children prefer the Likert scale.
In a study comparing various response options, reliability, validity, and feasibility are important outcomes. Feasibility was the main purpose of our study. It was extensively tested by assessing missing data, and asking children which options they found easiest to fill out and which options they preferred. Reliability is done by performing test-retest procedures. In test-retest studies, all circumstantial details should be identical at the first and second assessment. Because of our out-patient clinic setting with a mixture of relatively well and sick patients, we did not find it ethically justifiable to ask patients to return to the clinic just for research purposes. However, we find this would be an interesting study in the context of a clinical study or a mailing survey.

For the assessment of criterion validity, a gold standard is required. Only one study compared two response options in children with an objective gold standard as criterion, assessing cough frequency. The authors concluded the verbal categorical response option to show the highest correlation with objectively measured cough frequency. More often however, no gold standard is available. We therefore looked at concurrent validity, calculating correlation coefficients between the response options, assuming they would all somehow reflect the child’s true answer. The questions were therefore chosen to be simple, encompassing three types of questions frequently encountered in QoL instruments (affection, opinion and frequency). The types of questions however did not appear to influence the correlations in any way. Correlations were generally high, with the exception of the coefficients for the 6 and 7 year old children. Apparently, these children are not very consistent in answering on the various response options. The group of children was however small, prohibiting strong final conclusions.

As for feasibility, 3% of all questions with the simple VAS as answering option were not filled out. Apparently, these children found it hard to answer questions with this answering option. The Likert scale was understood more easily. This is reflected by the fact that all children regardless of age rated the Likert scale as easiest and they liked this option best. Therefore, the hypothesis of age related response patterns was not supported by the data.

It is interesting to see that there was a difference between immigrant and native children in ease and preference for the different response options, since the simple VAS was rated higher by the immigrant children. They also needed more time to fill in the questionnaires. It can be expected that this has to do with the fact that immigrant children might be less verbally equipped and therefore prefer the simple VAS. However, reliability of the VAS was reported to be low in illiterate people. It is however not completely justified to regard immigrant children as illiterate, since most attend regular Dutch schools in our country.

Although our data suggest very high correlation, we wonder whether this correlation was influenced by the lack of time between the different questionnaires. Also, since the three pages were combined in one package, children had the availability of seeing previous answers. In former studies comparing Likert scale and visual analogue scales, investigators introduced an interval between the different questionnaires, assuming that the first answer
would not be recalled when the second and the third scale were completed. On the contrary Downie et al. showed that correlation is maintained whether completion of the scales is separated in time or not. This may suggest that there is no significant carry-over from one scale to the next.

Pantel and Lewis identified a position bias in children’s responses on questionnaires, where they tended to choose the first answer among response options. This effect was strongest in younger children. Other authors stated that when using the simple VAS, children and elderly stick to the anchors without using the full scale range. On the other hand, in adults the VAS and Likert scales often encounter “end-aversion bias”, where people tend not to use the two extremes of the scale. We found children to stick to the first answering category on the simple VAS, which remained apparent when the categories were rearranged into uneven categories. In this way the advantage of the increase of answering precision in the simple VAS is lost. On the Likert scale more children use options two and three (a little bit, sometimes), suggesting end-aversion bias. This problem might be overcome by increasing the number of answering categories to 7 or even 9.

In conclusion, the simple VAS, numeric VAS and the Likert scale can be considered comparable in children 8 years and older. However, children of all ages prefer the Likert scale because of ease. The simple VAS, which we expected to be popular amongst children, scored low. We therefore recommend the use of Likert scaling options in questionnaires for children aged six years and older.

Figures 1a, 1b and 1c.
Response options to question: How do you like school?

1a.

- very nice
- nice
- in between
- annoying
- very annoying

1b.

very nice 10 9 8 7 6 5 4 3 2 1 very annoying

1c.

very nice _______________________________ very annoying

Format of scales: A = Likert scale, B = numeric VAS, C = simple VAS.
Chapter 3.4 Likert versus VAS

References
