RESEARCH

Open Access

How do children understand and respond to the EQ-5D-Y-3L? A mixed methods study in a community-based sample of 6–12-year-olds

Diana Khanna[1*] , Kiri Lay[1] , Jyoti Khadka[1,2] , Christine Mpundu-Kaambwa[3] , Julie Ratcliffe[1] and in collaboration with the Quality of Life in Kids: Key evidence to strengthen decisions in Australia (QUOKKA) project team

Abstract

Background The EQ-5D-Y-3L is widely used for measuring and valuing HRQoL in paediatric populations. This mixed methods study used the EQ-5D-Y-3L measure and applied a retrospective think-aloud approach to examine the self-report validity in children of varying chronological age.

Methods A mixed methods study was conducted in a community-based sample of 39 children aged 6–12 years. In a semi-structured interview, children self-completed the EQ-5D-Y-3L and then engaged in retrospective think-aloud. Conversations were audio-recorded and transcribed for analysis in NVivo using the Tourangeau four-stage response model framework to assess comprehension, judgment, recall, and response mapping issues. Fisher's exact test was used to assess the differences between child-self reported HRQoL across subgroups. The inter-rater agreement between child-parent dyads was assessed with CCC for overall HRQoL and Gwet's AC_1 for dimension level HRQoL.

Results Overall, response issues were detected in $n = 18$ (46%) children. Comprehension issues were apparent in the "having pain or discomfort" dimension where children found it challenging to understand 'discomfort'. Recall-related issues were observed where children's responses were influenced by their typical tendencies (e.g., being *usually* worried) or past incidences (e.g., feeling pain *sometimes*). Judgement-related issues were the most common, particularly in the "doing usual activities" dimension, where children tended to respond based on their *self-perceived ability* to engage in activities rather than health-related limitations. None of the participants were found to have problems with response mapping. A *healthy* lifestyle that included diet and exercise was a notable consideration in EQ VAS ratings.

The younger age groups had a higher proportion of response issues (6–7 years: 64%, 8–10 years: 62%), compared to older children (11–12 years: 20%). Moreover, children with response issues demonstrated significantly lower EQ-5D-Y-3L scores (mean = 0.78, se = 0.04) as compared to those without (mean = 0.95, se = 0.02) (*p*-value < 0.001). The overall inter-rater agreement was higher for those without any response issues (CCC = 0.33) than those with (CCC = 0.14). Additionally, higher agreement was noted across all the five dimensions in the subgroup with no response issues relative to those with.

*Correspondence:
Diana Khanna
diana.khanna@flinders.edu.au
Full list of author information is available at the end of the article

Khanna *et al. Health and Quality of Life Outcomes*   (2024) 22:105

Page 2 of 11

Conclusions Children in the general community may have different perceptions of HRQoL when responding to the EQ-5D-Y-3L possibly due to their limited experience with health-related challenges. The retrospective think-aloud approach adopted highlighted the relatively higher prevalence of response issues in the younger children (ages < 11 years), indicating the need for careful interpretation of self-reported HRQoL using the current version of the EQ-5D-Y-3L in this population.

Keywords Child self-report validity, Preference-based HRQoL measure, Retrospective think-aloud approach, Mixed methods study, Parent–child agreement

Background

Child health-related quality of life (HRQoL) is an important outcome for economic evaluation, quality assessment and epidemiological studies focused in child populations [1]. In a review by Chen and Ratcliffe, nine validated generic preference-based measures were identified for application in children and adolescents to measure and value HRQoL [2]. The use of child-specific HRQoL measures enables children to self-report subjective aspects of HRQoL that are unique to them [2]. One of the most widely applied child-specific preference-based measures with both self and proxy versions, is the EQ-5D-Y-3L, developed by the Euroqol group [2, 3]. The EQ-5D-Y-3L has been incorporated in a variety of economic evaluations alongside randomised control trials in Australia and internationally [4–7].

Proxy-reports of child HRQoL are useful when children are too young to self-report [8]. However, a review by Wolstenholme and colleagues found that the majority of studies applying preference-based measures used proxy-reports of child HRQoL, frequently obtained from parents, even for children up to 18 years old [9]. Despite these findings, there is a general consensus that children aged 8 and above are able to self-report their own HRQoL [10, 11]. Consistent with this, the use of EQ-5D-Y has been recommended in children aged 8–15 years [12]. According to the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Good Research Practices Patient-Reported Outcomes (PRO) task force report, the reliability of self-reporting in children aged 5–7 is unclear [8]. Nevertheless, Varni et al. have suggested that children aged 5 can self-report their HRQoL reliably and validly using an age-appropriate measure [13]. Moreover, for 6–7-year-olds, interviewer-administered measures have been reported to yield reliable results [14, 15].

Traditionally, less emphasis has been placed on how children interpret and respond to self-report dimensions within HRQoL measures, with greater attention given to establishing the psychometric properties of newly developed measures. Typically, when evaluating the psychometric properties of a measure, its validity, reliability, responsiveness, acceptability, feasibility, and consistency are assessed [3, 16, 17]. While these properties can provide critical insights, they may not offer a complete picture of the self-report reliability in young children. Furthermore, a wide range of studies, including those focussing on psychometric validation [18–21] and population health assessments [22–24], often include predominantly healthy children from the general population either as sole participants or as comparators. Therefore, to ensure that the child's self-reports are valid, it is essential to examine whether children, especially those from the general population, can understand the concepts conveyed by HRQoL dimensions and provide meaningful responses [25].

Bevans et al. suggest that cognitive interviews with children can assist in measuring their cognitive capacity and provide evidence of their understanding of HRQoL dimensions [26]. Cognitive interview methods such as the think-aloud, which involves articulating one's thoughts while engaging in a task (concurrent) or after task completion (retrospective), are key tools for assessing the quality of responses and for determining whether the HRQoL measure is generating the intended information from the respondents [27, 28]. Cremeens et al. conducted a qualitative study using think-aloud in the USA to investigate how children, aged 5–9 years, understood and responded to questions in the TedQL [29], a generic measure of quality of life for children aged 3–8 years [25]. Their findings suggested that older children (7-9 years) had a better understanding and interpretation of the TedQL items than younger children (5-6 years) [25]. While cognitive interviews have been conducted in other countries with children and adolescents to assess their understanding of HRQoL measures, specifically non-preference based [30, 31], there is limited qualitative research in Australia exploring how children comprehend and respond to child-specific preference-based measures for HRQoL.

A previous systematic review reported a low level of agreement between child-self and proxy reports of HRQoL using preference-based HRQoL measures [32]. However, the extent to which this discrepancy arises from difference in the understanding of HRQoL dimensions remains underexplored. To address this gap, the aims of this study were twofold: 1) to assess the self-report reliability of responses to the EQ-5D-Y-3L in a

community sample of school-aged children (6-12 years) and 2) to determine the impact of response issues on child self-reported HRQoL and inter-rater agreement between the child and a parent proxy assessor.

Methods

Study design

This study used a convergent mixed method design to gain a comprehensive understanding of child-self reporting of HRQoL by drawing on the strengths of both quantitative and qualitative methodologies with equal emphasis given to both [33]. Integration was achieved at the data analysis stage by using a data transformation model which utilises the quantified qualitative data in the analysis [33]. In addition, the study reported the qualitative findings to further elaborate on the integrated results.

Participants

An independent social research company facilitated participant recruitment for this cross-sectional study. An invitation letter with the details of the study was sent to eligible parent–child dyads. To achieve broad representation of the general population, the sample was selected using a stratified random sampling method that accounted for socio-demographic characteristics, and common health conditions amongst child populations including asthma, anxiety disorders, conduct disorders, depressive disorders, autism spectrum disorders, and dental caries [34]. Children aged 6–12 years who were able to read and understand English were eligible for the study and were invited, along with their parents, to participate in a semi-structured face-to-face interview in April 2021. The interviews were conducted with a researcher at Flinders University, South Australia.

Ethics

Each participating parent provided written informed consent, and verbal assent was obtained from the child. This study was conducted in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. The study was approved by the Flinders University's Human Research Ethics Committee (Project ID 4178).

Measures

The study utilised the EQ-5D-Y-3L, a version of the EQ-5D-3L specifically adapted for children aged 8 to 15, as well as its Proxy version 1 (rate child's HRQoL from proxies' perspective) to assess the child's HRQoL and examine the inter-rater agreement between self and proxy reporting of HRQoL [35]. There are five dimensions within both versions, including "walking about", "looking after myself", "doing usual activities", "having pain or discomfort", and "feeling worried, sad or unhappy" with the recall period limited to 'today'. Respondents can indicate severity on any of the three levels of problems (no problems, some problems, a lot of problems) for each dimension. Both, the EQ-5D-Y-3L and its Proxy version 1, also include a visual analogue scale (EQ VAS) allowing respondents to rate the child's overall health status on a scale from 0 to 100, with 0 indicating the worst imaginable health and 100 the best imaginable health.

Since there is currently no value set available for the EQ-5D-Y-3L in Australia, the Australian adult general population value set for the EQ-5D-3L was applied [36]. This value set was applied to both proxy and self HRQoL ratings. A previous study in a larger sample conducted a sensitivity analysis using the German value sets for the EQ-5D-Y-3L against the Australian and found no significant difference in both self and proxy reported outcomes [37].

Data collection

Children completed the self-report version of the EQ-5D-Y-3L and a single self-rated general health (SRH) item question ("In general, would you say your health is poor, fair, good, very good, or excellent?" [38]) on a laptop computer via the REDCap (Research Electronic Data Capture) software. REDCap is a secure, web-based application designed to support data capture and management for research studies [39, 40]. The on-screen process of children completing the measure was recorded in a video (screen recording) to later replay in the qualitative stage of the interview. Simultaneously, the parent was asked to complete a hard copy of EQ-5D-Y-3L (Proxy version 1). The parents were also asked to report on the child's age, gender and long-term health condition/s if any.

Following the completion of the questionnaires, face-to-face retrospective think-aloud interviews were conducted using an age-appropriate interview guide specifically developed for this study. The retrospective think-aloud method was chosen as it allows uninterrupted completion of the questionnaire by the child and reduces the workload during the task [41, 42]. The interviewer asked the children to retrospectively think-aloud whilst parents wore noise cancelling headphones such that their responses were not unduly influenced by any conversations taking place between the interviewer and the child. The recorded video of the child completing the EQ-5D-Y-3L on the laptop screen was played back to them to assist with the process of think-aloud and prompt recollection of their responses. This conversation was recorded on a voice recorder.

Data analysis

Qualitative data analysis

Interviews were transcribed verbatim, and the data was analysed using NVivo qualitative analysis software (release 1.3). The Tourangeau four-stage response model framework identified response issues at each of the four stages of survey response, comprehension, recall, judgement, and response mapping (Table 1) [43]. This framework was successfully utilised with a think-aloud protocol by members of our team to assess the self-report reliability of the EQ-5D-5L in older adults with varying levels of cognitive impairment and dementia [44, 45].

Two coders (DK and KL) independently coded the transcripts and identified response issues in accordance with the Tourangeau response model framework. The identified issues were then categorised into one or more type of response issues, with potential overlap, depending on the stage at which the response process was deemed erroneous. Codes were created based on the perceived source of response issue within these categories. The EQ VAS was not coded for response issues under the framework; however, an inductive thematic analysis was undertaken. Open coding was conducted followed by the iterative development and refinement of themes to understand perceptions and interpretations of the EQ VAS [46]. During the analysis, the codes and findings were thoroughly discussed, and any discrepancies were addressed through discussion by the research team until a consensus was reached.

Quantitative data analysis

Qualitative data, comprising response issues using the Tourangeau four-stage response model framework were integrated into the quantitative data analysis. Data transformation was achieved by converting the response issues into numerical variables that indicated the type of issue, as detailed in the qualitative analysis section. The sample was described using descriptive statistics based on the response issues of comprehension, judgment, recall, and response mapping for the subgroups: 1) age-group: 6–7 years old, 8–10 years old or 11-age group, 2) presence of long-term condition: yes or no, and 3) gender: female or male. A dichotomous variable representing the presence of response issue was used.

Data was analysed using Stata 16.1 employing statistical methods appropriate for analysis of non-normal data [47]. To evaluate the agreement between raters, the concordance correlation coefficient (CCC) was used for overall HRQoL [48] due to the skewed sample distribution [49]. Dimension-level HRQoL was analysed using Gwet's agreement coefficient (AC_1). Both CCC and Gwet's AC_1 range between −1 and 1. To assess the strength of agreement, Altman's scale was used which categorises the magnitude of the coefficient as poor (≤ 0.2), fair (0.21 to 0.4), moderate (0.41 to 0.6), good (0.61 to 0.8), or very good (> 0.8) [50, 51]. Fisher's exact test was used to explore statistical differences between child-self reported HRQoL across subgroups. In this study, the statistical significance level was set at 0.05. The socio-economic disadvantage of the postal area was assessed using the Socio-Economic Indexes for Australia (SEIFA) area-decile number. The first six decile numbers were categorised as disadvantaged quintiles (quintiles 1 to 3), while the last four decile numbers were classified as advantaged quintiles (quintiles 4 and 5) [52].

Table 1 Tourangeau four-stage response model framework

Stage	Description
Comprehension	Involves problems with the interpretation of the HRQoL dimensions, where participants encounter problems understanding or have misunderstandings of words, phrases, or response options. Issues may arise due to ambiguous wording or unfamiliar terms
Recall	Involves problems with the retrieval process, encompassing challenges such as using an incorrect time frame for information retrieval and difficulties in recalling relevant information that aligns with the measure's specified time frame
Judgment	Involves problems with assessing the information retrieved, including drawing upon irrelevant information or inadequate evaluation, potentially leading to under or over-reporting of health states. This stage is often closely linked to comprehension, as it entails the evaluation of recalled information to determine its relevance to the specific question and is influenced by how the question was understood in the first place
Response Mapping	Addresses issues with the response categories, such as dissatisfaction with the provided options or inappropriate application of them (e.g., selecting multiple options when only one is permissible, or reluctance to choose any). It also includes cases where the stated answer (verbal protocol data) is misaligned with the chosen answer (survey response data), reflecting a disconnection in the final stage of the survey

(Adapted from Lay, K. et al. [45])

Khanna *et al. Health and Quality of Life Outcomes*   (2024) 22:105

Page 5 of 11

Table 2 Sociodemographic characteristics of the child and parent participants

	N = 39	% sample
Child age		
Mean (standard deviation)	9.1 (2)	
Median (IQR)	9 (4)	
Parent age[a]		
Mean (standard deviation)	41.1 (5)	
Median (IQR)	40 (8)	
Child gender		
Female	22	56
Male	17	44
Parent gender[a]		
Female	31	82
Male	7	18
Child self-rated general health[b]		
Excellent	6	16
Very good	16	43
Good	12	33
Fair	3	8
Child long term health condition		
Yes	28	72
No	11	28
Specific health condition		
Autism Spectrum Disorder	2	18
Asthma	5	45
Dental caries	2	18
Sleep disorders	2	2
Socio-economic condition according to post-code[b]		
Relatively advantaged quintile (SEIFA decile 7,8,9,10)	27	73
Relatively disadvantaged quintile (SEIFA decile 1,2,3,4,5,6)	10	27

SEIFA Socio-Economic indexes for Australia

[a] n = 38 reported

[b] n = 37 reported

Results
Integrated findings
Sample characteristics

Table 2 presents the sociodemographic characteristics of the child-parent dyads. A total of 39 children participated in the interview, with a median age of 9 (IQR = 4) years. Females were slightly over-represented, accounting for 56% of the participants. Most children reported themselves to be in 'good' to 'excellent' health, with only 8% reporting 'fair' health on the single SRH item. According to parental reports, 28% of the children had one of the following conditions: asthma (45%), autism spectrum disorder (18%), dental caries (18%), and sleep problems (18%). The parents of the children in the sample had a median age of 41 (IQR = 9), and approximately one-fifth of the dyads consisted of father-child pairs. When considering SEIFA area-decile numbers, a lower proportion of respondents resided in postcodes associated with relatively disadvantaged quintiles (25%) compared to the Australian population.

Response issues by age-group, presence of long-term health condition and gender

A total of 46% (*n* = 18) children experienced at least one or more response issues. The highest proportion of response issues was observed among children in the youngest age-group (6–7 years: 64%), followed by the 8–10 year-olds (62%), while the oldest children (11–12 years) had the lowest proportion of response issues (20%). As illustrated in Table 3, the most common response issues were related to judgement (28%) and recall (23%), while comprehension issues were relatively less frequent (18%). Comprehension issues were mainly observed among the youngest age-group, while judgement and recall issues were predominant in children aged 8–10 years.

Table 3 Description of response issues by age-group, presence of long-term health condition and gender

Demographic characteristic	Comprehension n (%)	Judgement n (%)	Recall n (%)	Total (unique) n (%)
Age group				
6–7 years (*N* = 11)	5 (0.45)	3 (0.27)	3 (0.27)	7 (0.64)
8–10 years (*N* = 13)	2 (0.15)	7 (0.54)	4 (0.31)	8 (0.62)
11–12 years (*N* = 15)	0	1 (0.07)	2 (0.13)	3 (0.2)
Presence of long-term health condition				
No (*N* = 28)	4 (0.14)	7 (0.25)	5 (0.18)	12 (0.43)
Yes (*N* = 11)	3 (0.27)	4 (0.36)	4 (0.36)	6 (0.55)
Gender				
Female (*N* = 22)	2 (0.09)	8 (0.36)	7 (0.32)	10 (0.44)
Male (*N* = 17)	5 (0.29)	3 (0.18)	2 (0.12)	8 (0.47)
Total (*N* = 39)	7 (0.18)	11 (0.28)	9 (0.23)	18 (0.46)

Overall, there was a potential association between age group and response issues albeit the significance was only marginal (*p*-value = 0.08; Fisher's exact test). However, significant differences were observed between age-group and two specific types of response issues, namely comprehension (*p*-value = 0.02) and judgment (*p*-value = 0.03).

Children with reported health conditions and male children exhibited a marginally higher proportion of response issues compared to their respective comparison groups, as shown in Table 3. However, these differences were not statistically significant.

Relationship between response issues and 1) self-reported HRQoL scores and 2) inter-rater agreement

Children with response issues had significantly lower (*p*-value < 0.001) EQ-5D-Y-3L scores (mean = 0.78, se = 0.04) compared to those with no response issues (mean = 0.95, se = 0.02). Table 4 shows the inter-rater agreement for the dimension level and overall EQ-5D-Y-3L scores according to the presence of response issues. Poor child-parent agreement (CCC = 0.14, 95% CI = −0.31, 0.54) was observed among children with response issues, while a higher agreement (CCC = 0.33, 95% CI = −0.06, 0.63) was observed in children with no response issues.

Agreement at the dimension level was lower for all dimensions in children with response issues, except for the "feeling worried, sad or unhappy" dimension. The lowest agreement was observed within the "doing usual activities" and "having pain/discomfort" dimensions. Correspondingly, the highest number of problems reported (Table 5) as well as response issues was observed within the "doing usual activities" dimension with judgement and "having pain/discomfort" dimension with comprehension (Fig. 1).

Response issues by EQ-5D-Y-3L dimensions and EQ VAS-qualitative findings

Table 6 highlights the key qualitative findings related to response issues encountered by children in understanding and responding to the EQ-5D-Y-3L and the EQ VAS.

In the "walking about" dimension, the response issue was due to conflation of health-related ability to walk with ability to, for instance, walk without tripping or unaccompanied by adults. In the "looking after myself" dimension, children reported non-health related issues such as frustration with clothes or dislike of washing themselves. The responses were based on their overall patterns of behaviour rather than issues on the day of assessment. In the "doing usual activities" dimension, all judgement related response issues were general non-health related inabilities rather than health related limitations. Similar to other dimensions, recall related

Table 4 Inter-rater agreement for dimension level and overall EQ-5D-Y-3L scores by presence of response issues

	Walking about	Looking after myself	Doing usual activities	Having pain/ discomfort	Feeling worried, sad or unhappy	Overall
	AC$_1$ (95% CI)	AC$_1$ (95% CI)	AC$_1$ (95% CI)	AC$_1$ (95% CI)	AC$_1$ (95% CI)	CCC (95% CI)
Overall sample (*N* = 38)	0.83 (0.69, 0.97)	0.74 (0.56, 0.91)	0.63 (0.43, 0.84)	0.61 (0.41, 0.82)	0.57 (0.35, 0.78)	0.28 (−0.03, 0.54)
No response issue (*N* = 21)	0.9 (0.75, 1)	0.85 (0.66, 1)	0.85 (0.66, 1)	0.73 (0.49, 0.98)	0.53 (0.23, 0.84)	0.33 (−0.06, 0.63)
Some response issue (*N* = 17)	0.74 (0.46, 1)	0.59 (0.25, 0.92)	0.33 (−0.07, 0.72)	0.45 (0.09, 0.82)	0.61 (0.29, 0.94)	0.14 (−0.31, 0.54)

AC1 Gwet's Agreement Coefficient, *CCC* Concordance Correlation Coefficient, Altman's scale interpretation: Less than or equal to 0.2 = Poor, between 0.21 & 0.4 = Fair, between 0.41 & 0.6 = Moderate, between 0.61 & 0.8 = Good, between 0.81 & 1 = Very Good

Table 5 Frequency of problems reported with the child self-report version of the EQ-5D-Y-3L

Dimensions	No problems n (%)	Some problems n (%)	A lot of problems n (%)
Walking about	33 (0.87)	5 (0.13)	
Looking after myself	33 (0.87)	5 (0.13)	
Doing usual activities	29 (0.76)	9 (0.24)	
Having pain/discomfort	27 (0.71)	10 (0.26)	1 (0.03)
Feeling worried, sad or unhappy	32 (0.84)	6 (0.16)	
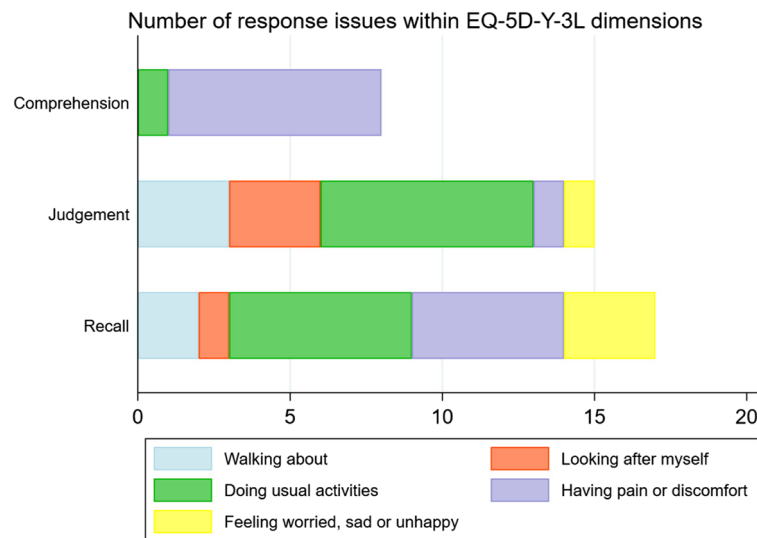
Fig. 1 Response issues within EQ-5D-Y-3L dimensions

Table 6 Selected quotes from the qualitative analysis of response issues by EQ-5D-Y-3L Dimensions, Participant Age, and Selected Responses

Quote	Dimension	Age of Participant	Selected Response
"because, sometimes when I walk, it hurts, over here sometimes"	Walking About	8 yrs	Some problems
"Sometimes I make like mistakes in like – sometimes in sports games I make a mistake about tripping or accidentally hurting by bumping because in – we play basketball sometimes and once we were playing and I accidentally bumped into someone and made them fall and I helped them up though"	Doing Usual Activities	9 yrs	Some problems
"Because I sometimes when – like it's really hard to figure stuff out… I meant like when I'm doing maths and stuff"	Doing Usual Activities	8 yrs	Some problems
"I don't have pain or discomfort, because I usually always fit in, and no-one forces me to not do that stuff, like be rude to me…"	Having Pain or Discomfort	7 yrs	No problems
"Discomfort, I thought it means you're not comfortable with all your friends talking behind your back. That's what I thought it meant"	Having Pain or Discomfort	8 yrs	No problems
"Well, I'm not like – no-one's really happy all the time. But I wouldn't say that I'm sad all the time, I'm just at a normal, maybe some things might have made me upset, like small petty things"	Feeling Worried, Sad, or Unhappy	11 yrs	Some problems
"Yeah. I'm not like the happiest, but I'm not upset – just like pretty good, I'm feeling"	EQ VAS	11 yrs	Rating: 71
"Like I could be healthier with my choices of eating and stuff, and I could do – eat more healthier things and do more healthier things. But I'm not really unhealthy. I don't eat heaps of chocolate and not much vegetables and fruit. I still do quite a few – I do basketball for a few hours a week and stuff. I still do exercising and stuff"	EQ VAS	11 yrs	Rating: 75

problems were identified when children responded based on their general behavioural tendencies or specific events outside of the recall period.

Comprehension issues were predominant within the "having pain or discomfort" dimension wherein there was a lack of understanding of 'discomfort' or 'pain and discomfort' was interpreted to mean 'emotional pain'. The judgement related response issue within this dimension was related to the concept of health. For example, one child said they did not have 'pain or discomfort' as they 'work on their health'. Only a few response issues were identified within the "feeling worried, sad or unhappy"

Khanna *et al. Health and Quality of Life Outcomes*   (2024) 22:105

Page 8 of 11

dimension, mostly related to recall (e.g., recalled their inclination to worry).

A common theme emerging from the EQ VAS ratings was the association of the concept of health with emotional wellbeing and lifestyle. Children considered their emotional health or healthy habits such as eating fruits and vegetables and exercising when completing the EQ VAS.

Discussion

Despite its ubiquitous use, research exploring how children understand and interpret the items in the EQ-5D-Y is limited [22, 53–58]. This study aimed to address this gap in the literature by investigating the response process in children aged 6–12 years who self-completed the EQ-5D-Y-3L. Using the Tourangeau four-stage response model, the findings indicated that the response process varied depending on the age group, with children ≤ 10 years of age demonstrating a higher proportion of comprehension, judgement, and recall related issues as compared to older children.

Wille et al. conducted a multinational study to adapt the EQ-5D-Y-3L from the adult EQ-5D-3L [59]. They found that the items were well accepted and generally comprehensible in a predominantly healthy sample of children and adolescents aged 8–18 years. Nevertheless, the major challenge identified by the study was children's difficulty distinguishing between health-related impairment and age-related inability [10]. The findings of a South African study by Amien and colleagues also indicated younger children (ages 5–7 years) reported a significantly higher frequency of problems with "looking after myself" as compared to older children (ages 8–10 years), primarily due to their need for assistance, which was unrelated to a medical condition [60]. In this study, only two children, ages 8 and 11 years, respectively, reported problems with "walking about" and "looking after myself" due to age-related independence/limitation. However, in a similar vein, a predominant issue identified in this study was judgement related, particularly for "doing usual activities". Children responded thinking about their non-health related ability to perform the task, rather than their health-related limitations. More specifically, they reported problems if they considered their general limitations in relation to athletic or academic performance in comparison to other children. This is consistent with the findings of Cremeens et al., who reported that children frequently used social comparisons to judge their own ability to perform certain tasks [25].

Children also considered their emotional well-being when responding to "having pain and discomfort" and their overall health using the EQ VAS. Children tended to associate the word 'discomfort' with emotional discomfort. Amien et al. reported that children may not be familiar with the word "discomfort"; however, in their study, this lack of familiarity did not affect their understanding of the question being asked [60]. A previous study in the USA found that school-age children (aged 5–11 years) associated the term "healthy" with behaviours such as eating fruits and vegetables [11]. In this study, children also considered a "healthy" lifestyle, which included diet and exercise, when evaluating their overall health using the EQ VAS. They also considered their level of happiness, an indicator of quality of life [61], when evaluating their overall health.

The EQ-5D-Y requires the child to report their health status 'today'. In their review, Arbuckle et. al state that children under the age of 6 years may have limited introspective abilities and struggle with distinguishing between past, present, and future [62]. In this sample, while many children did understand that they had to report their health status 'today', recall issues were identified mainly with "having pain/discomfort". Specifically, it was observed that some children based their responses on past experiences of pain or discomfort, even if they were not experiencing any pain or discomfort the day of the interview. Future investigations comparing child specific HRQoL measures with different timeframes (e.g., EQ-5D-Y and the PedsQL™) considering children's developmental capabilities and potential recall biases may be needed to provide deeper insights into children's responses to HRQoL assessments.

Most importantly, this study found that the presence of response issues was associated with lower self-reported HRQoL scores. Whilst child self-reported HRQoL did not vary by the presence or absence of long-term health-condition/s, it was found to vary by response issues; children with identified response issues reported a lower HRQoL overall compared to the sub-group with no response issues (mean difference = 0.17, se = 0.04). It is possible that this lower reported HRQoL may result from incorrect interpretation and response to the HRQoL dimension/s themselves, rather than actual health-related limitations. Similarly, dimension level child-parent agreement was lower in the sub-group with identified response issues (relative to the children with no identified response issues) for all dimensions, except "feeling worried, sad or unhappy". The higher level of agreement within this dimension may be attributed to the reason for the response issues. Meaning, children reported on their general tendency to worry which likely would have also been the same reasoning guiding the parent's proxy response resulting in a higher level of agreement.

The main limitation of this study was that the children in the sample were from the general population and may not have had experience or understanding of the impact

of health conditions on the dimensions included in the measures used. However, this cohort is often used in research-related (e.g., studies relating to psychometric validation of measures validation [18–21] and population health assessments [22–24]), making these findings particularly relevant. Prior studies have investigated the reliability and validity of self-reports in children as young as 5 years of age using only quantitative analysis [13–15, 63]. However, this study provides valuable insights into the reliability of self-reported EQ-5D-Y-3L responses and inter-rater agreement in Australian children using a mixed methods study design.

Conclusion

This study aimed to examine the validity of children's responses when self-reporting their HRQoL using the EQ-5D-Y-3L and its impact on the HRQoL scores and inter-rater agreement. The findings highlighted that younger children (particularly children aged 6–7 years and 8–10 years) may face several challenges in their understanding and comprehension of the HRQoL dimensions relative to older children (aged 11–12 years). These challenges may result in younger children self-reporting a lower level of HRQoL and inconsistencies with parental proxy reported HRQoL, which are not solely based on the actual quantification of health-related impairments intended to be captured by the measure.

Abbreviations

AC	Agreement Coefficient
CCC	Concordance Correlation Coefficient
CI	Confidence Intervals
HRQoL	Health-Related Quality of Life
QALY	Quality Adjusted Life Year
USA	Unites States of America
VAS	Visual Analogue Scale

Declarations

Author details
[1]Health and Social Care Economics Group, Caring Futures Institute, College of Nursing and Health Sciences, Flinders University, Adelaide 5042, Australia. [2]Registry of Senior Australians, Healthy Ageing Research Consortium, South Australian Health and Medical Research Institute, Adelaide, Australia. [3]Melbourne School of Population and Global Health, University of Melbourne, Victoria, Australia.

References

1. Brazier J, Ratcliffe J, Saloman J, Tsuchiya A. Measuring and valuing health benefits for economic evaluation. OXFORD university press. 2017.
2. Chen G, Ratcliffe J. A Review of the Development and Application of Generic Multi-Attribute Utility Instruments for Paediatric Populations. Pharmacoeconomics. 2015;33(10):1013–28. https://doi.org/10.1007/s40273-015-0286-7.
3. Golicki D, Młyńczak K. Measurement Properties of the EQ-5D-Y: A Systematic Review. Value in Health. 2022;25(11):1910–21. https://doi.org/10.1016/j.jval.2022.05.013.
4. Wang HI, Wright B, Tindall L, Cooper C, Biggs K, Lee E, et al. Cost and effectiveness of one session treatment (OST) for children and young people with specific phobias compared to multi-session cognitive behavioural therapy (CBT): results from a randomised controlled trial. BMC Psychiatry. 2022;22(1):547. https://doi.org/10.1186/s12888-022-04192-8.
5. Wang HI, Wright BD, Bursnall M, Cooper C, Kingsley E, Le Couteur A, et al. Cost-utility analysis of LEGO based therapy for school children and young people with autism spectrum disorder: results from a randomised controlled trial. BMJ Open. 2022;12(1): e056347. https://doi.org/10.1136/bmjopen-2021-056347.
6. Mihalopoulos C, Chen G, Scott JG, Bucholc J, Allen C, Coghill D, et al. Assessing Outcomes for Cost-Utility Analysis in Children and Adolescents With Mental Health Problems: Are Multiattribute Utility Instruments Fit for Purpose? Value Health. 2023;26(5):733–41. https://doi.org/10.1016/j.jval.2022.12.007.
7. Rombach I, Wang K, Tonner S, Grabey J, Harnden A, Wolstenholme J. Quality of life, healthcare use and costs in "at-risk" children after early antibiotic treatment versus placebo for influenza-like illness: within-trial descriptive economic analyses of the ARCHIE randomised controlled trial. BMJ Open. 2022;12(4): e049373. https://doi.org/10.1136/bmjopen-2021-049373.
8. Matza LS, Patrick DL, Riley AW, Alexander JJ, Rajmil L, Pleil AM, et al. Pediatric patient-reported outcome instruments for research to support medical product labeling: report of the ISPOR PRO good research practices for the assessment of children and adolescents task force. Value in Health. 2013;16(4):461–79. https://doi.org/10.1016/j.jval.2013.04.004.
9. Wolstenholme JL, Bargo D, Wang K, Harnden A, Räisänen U, Abel L. Preference-based measures to obtain health state utility values for use in economic evaluations with child-based populations: a review and UK-based focus group assessment of patient and parent choices. Qual Life Res. 2018;27(7):1769–80. https://doi.org/10.1007/s11136-018-1831-6.
10. Wille N, Badia X, Bonsel G, Burström K, Cavrini G, Devlin N, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. Qual Life Res. 2010;19(6):875–86. https://doi.org/10.1007/s11136-010-9648-y.

11. Rebok G, Riley A, Forrest C, Starfield B, Green B, Robertson J, et al. Elementary school-aged children's reports of their health: A cognitive interviewing study. Qual Life Res. 2001;10(1):59–70. https://doi.org/10.1023/A:1016693417166.

12. Foundation E. R. EQ-5D-Y User Guide. 2020.

13. Varni JW, Limbers CA, Burwinkle TM. How young can children reliably and validly self-report their health-related quality of life?: an analysis of 8,591 children across age subgroups with the PedsQL 4.0 Generic Core Scales. Health Qual Life Outcomes. 2007;5:1–1. https://doi.org/10.1186/1477-7525-5-1.

14. Canaway AG, Frew EJ. Measuring preference-based quality of life in children aged 6–7 years: a comparison of the performance of the CHU-9D and EQ-5D-Y—the WAVES Pilot Study. Qual Life Res. 2013;22(1):173–83. https://doi.org/10.1007/s11136-012-0119-5.

15. Conijn JM, Smits N, Hartman EE. Determining at What Age Children Provide Sound Self-Reports: An Illustration of the Validity-Index Approach. Assessment. 2019;27(7):1604–18. https://doi.org/10.1177/1073191119832655.

16. Janssens A, Thompson CJ, Rogers M, Allen K, Green C, Jenkinson C, et al. A systematic review of generic multidimensional patient-reported outcome measures for children, part I: descriptive characteristics. Value in Health. 2015;18(2):315–33. https://doi.org/10.1016/j.jval.2014.12.006.

17. Jones R, Mulhern B, McGregor K, Yip S, O'Loughlin R, Devlin N, et al. Psychometric Performance of HRQoL Measures: An Australian Paediatric Multi-Instrument Comparison Study Protocol (P-MIC). Children (Basel, Switzerland), 2021;8(8). https://doi.org/10.3390/children8080714.

18. Ngwira LG, Maheswaran H, Verstraete J, Petrou S, Niessen L, Smith SC. Psychometric performance of the Chichewa versions of the EQ-5D-Y-3L and EQ-5D-Y-5L among healthy and sick children and adolescents in Malawi. Journal of Patient-Reported Outcomes. 2023;7(1):22. https://doi.org/10.1186/s41687-023-00560-4.

19. Verstraete J, Amien R, Scott D. Comparing Measurement Properties of the English EQ-5D-Y 3-Level Version With the 5-Level Version in South Africa. Value in Health Regional Issues. 2022;30:140–7. https://doi.org/10.1016/j.vhri.2021.12.004.

20. Scott D, Ferguson GD, Jelsma J. The use of the EQ-5D-Y health related quality of life outcome measure in children in the Western Cape, South Africa: psychometric properties, feasibility and usefulness - a longitudinal, analytical study. Health Qual Life Outcomes. 2017;15(1):12. https://doi.org/10.1186/s12955-017-0590-3.

21. Wu XY, Ohinmaa A, Johnson JA, Veugelers PJ. Assessment of children's own health status using visual analogue scale and descriptive system of the EQ-5D-Y: linkage between two systems. Qual Life Res. 2014;23(2):393–402. https://doi.org/10.1007/s11136-013-0479-5.

22. Wu X, Veugelers PJ, Ohinmaa A. Health Behavior, Health-Related Quality of Life, and Mental Health Among Canadian Children: A Population-Based Cohort Study. Front Nutr. 2021;8: 638259. https://doi.org/10.3389/fnut.2021.638259.

23. Åström M, Persson C, Lindén-Boström M, Rolfson O, Burström K. Population health status based on the EQ-5D-Y-3L among adolescents in Sweden: Results by sociodemographic factors and self-reported comorbidity. Qual Life Res. 2018;27(11):2859–71. https://doi.org/10.1007/s11136-018-1985-2.

24. Palacios-Cartagena RP, Adsuar JC, Hernández-Mocholí M, Carlos-Vivas J, Barrios-Fernández S, Garcia-Gordillo MA, et al. Health-Related Quality of Life Norm Data of the Peruvian Adolescents: Results Using the EQ-5D-Y. Int J Environ Res Public Health. 2021;18(16). https://doi.org/10.3390/ijerph18168735.

25. Cremeens J, Eiser C, Blades M. A qualitative investigation of school-aged children's answers to items from a generic quality of life measure. Child. 2007;33(1):83–9. https://doi.org/10.1111/j.1365-2214.2006.00665.x.

26. Bevans KB, Ahuvia IL, Hallock TM, Mendonca R, Roth S, Forrest CB, et al. Investigating child self-report capacity: a systematic review and utility analysis. Qual Life Res. 2020;29(5):1147–58. https://doi.org/10.1007/s11136-019-02387-3.

27. Padilla J-L, Leighton JP. Cognitive Interviewing and Think Aloud Methods. In: Zumbo BD, Hubley AM, editors. Understanding and Investigating Response Processes in Validation Research. Cham: Springer International Publishing; 2017. p. 211–28.

28. Ericsson KA, Simon HA. Verbal reports as data. Psychol Rev. 1980;87:215–51. https://doi.org/10.1037/0033-295X.87.3.215.

29. Lawford J, Volavka N, Eiser C. A generic measure of Quality of Life for children aged 3–8 years: results of two preliminary studies. Pediatr Rehabil. 2001;4(4):197–207. https://doi.org/10.1080/13638490210124033.

30. Hunter LC, Molitor F, Chafetz RS, Jane MM, Vogel LC, Betz R, et al. Development and Pilot Test of the Shriners Pediatric Instrument for Neuromuscular Scoliosis (SPNS): A Quality of Life Questionnaire for Children With Spinal Cord Injuries. J Spinal Cord Med. 2007;30(sup1):S150–7. https://doi.org/10.1080/10790268.2007.11754594.

31. Eddy L, Khastou L, Cook KF, Amtmann D. Item Selection in Self-Report Measures for Children and Adolescents with Disabilities: Lessons from Cognitive Interviews. J Pediatr Nurs. 2011;26(6):559–65. https://doi.org/10.1016/j.pedn.2010.07.004.

32. Khanna D, Khadka J, Mpundu-Kaambwa C, Lay K, Russo R, Ratcliffe J. Are We Agreed? Self- Versus Proxy-Reporting of Paediatric Health-Related Quality of Life (HRQoL) Using Generic Preference-Based Measures: A Systematic Review and Meta-Analysis. Pharmacoeconomics. 2022;40(11):1043–67. https://doi.org/10.1007/s40273-022-01177-z.

33. Creswell JW, Clark VLP. Designing and conducting mixed methods research. Sage publications. 2017.

34. Australian Institute of Health and Welfare. Health of children. Canberra: AIHW. 2020. Available from: https://www.aihw.gov.au/reports/australias-health/health-of-children.

35. Ravens-Sieberer U, Wille N, Badia X, Bonsel G, Burström K, Cavrini G, et al. Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study. Qual Life Res. 2010;19(6):887–97. https://doi.org/10.1007/s11136-010-9649-x.

36. Viney R, Norman R, King MT, Cronin P, Street DJ, Knox S, et al. Time trade-off derived EQ-5D weights for Australia. Value in Health. 2011;14(6):928–36. https://doi.org/10.1016/j.jval.2011.04.009.

37. Khanna D, Khadka J, Mpundu-Kaambwa C, Chen G, Dalziel K, Devlin N, et al. An Investigation of Inter-Rater and Intra-Proxy Agreement in Measuring Quality of Life of Children in the Community Using the EQ-5D-Y-3L. Pharmacoeconomics. 2024. https://doi.org/10.1007/s40273-024-01356-0.

38. Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. J Health Soc Behav. 1997;38(1):21–37.

39. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009;42(2):377–81. https://doi.org/10.1016/j.jbi.2008.08.010.

40. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform. 2019;95: 103208. https://doi.org/10.1016/j.jbi.2019.103208.

41. van den Haak M, De Jong M, Jan Schellens P. Retrospective vs. concurrent think-aloud protocols: Testing the usability of an online library catalogue. Behav Inform Technol. 2003;22(5):339–51. https://doi.org/10.1080/0044929031000.

42. Peute LW, de Keizer NF, Jaspers MW. The value of Retrospective and Concurrent Think Aloud in formative usability testing of a physician data query tool. J Biomed Inform. 2015;55:1–10. https://doi.org/10.1016/j.jbi.2015.02.006.

43. Tourangeau R, Rips LJ, Rasinski K. The Psychology of Survey Response. Cambridge: Cambridge University Press; 2000.

44. Engel L, Bucholc J, Mihalopoulos C, Mulhern B, Ratcliffe J, Yates M, et al. A qualitative exploration of the content and face validity of preference-based measures within the context of dementia. Health Qual Life Outcomes. 2020;18(1):178. https://doi.org/10.1186/s12955-020-01425-w.

45. Lay K, Crocker M, Engel L, Ratcliffe J, Charlton S, Hutchinson C. How do older adults receiving aged care services understand and respond to the EQ-5D-5L? A think-aloud study in residential care. Qual Life Res. 2023. https://doi.org/10.1007/s11136-023-03466-2.

46. Nowell LS, Norris JM, White DE, Moules NJ. Thematic Analysis: Striving to Meet the Trustworthiness Criteria. Int J Qual Methods. 2017;16(1):1609406917733847. https://doi.org/10.1177/1609406917733847.

47. StataCorp. Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC; 2019.

48. Chen C-C, Barnhart HX. Comparison of ICC and CCC for assessing agreement for data without and with replications. Comput Stat Data Anal. 2008;53(2):554–64. https://doi.org/10.1016/j.csda.2008.09.026.

Khanna *et al. Health and Quality of Life Outcomes* (2024) 22:105

Page 11 of 11

49. Khanna D, Khadka J, Mpundu-Kaambwa C, Ratcliffe J, in Collaboration with the Quality of Life in Kids: Key Evidence to Strengthen Decisions in Australia Project T. Child–Parent Agreement in the Assessment of Health-Related Quality of Life Using the CHU9D and the PedsQLTM. Applied Health Economics and Health Policy, 2023. https://doi.org/10.1007/s40258-023-00831-7.

50. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. BMC Med Res Methodol. 2013;13:61. https://doi.org/10.1186/1471-2288-13-61.

51. Klein D. Implementing a general framework for assessing interrater agreement in Stata. Stata Journal. 2018;18(4):871–901.

52. Australian Bureau Of Statistics. Socio-Economic Indexes for Areas (SEIFA) 2033.0.55.001 Canberra; 2018. Available from: https://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/756EE3DBEFA869EFCA258259000BA746/$File/SEIFA%202016%20Technical%20Paper.pdf.

53. Ryan JM, McKay E, Anokye N, Noorkoiv M, Theis N, Lavelle G. Comparison of the CHU-9D and the EQ-5D-Y instruments in children and young people with cerebral palsy: a cross-sectional study. BMJ Open. 2020;10(9):e037089. https://doi.org/10.1136/bmjopen-2020-037089.

54. Victoria-Montesinos D, Tárraga-Marcos A, Brazo-Sayavera J, Jiménez-López E, Gutiérrez-Espinoza H, Panisello Royo JM, et al. Adherence to the Mediterranean Diet and Health-Related Quality of Life during the COVID-19 Lockdown: A Cross-Sectional Study including Preschoolers, Children, and Adolescents from Brazil and Spain. Nutrients. 2023;15(3). https://doi.org/10.3390/nu15030677.

55. Garcia MC, Taylor-Haas JA, Rauh MJ, Toland MD, Bazett-Jones DM. Sport Specialization in Middle School and High School Long-Distance Runners. J Athl Train. 2021;56(9):1003–9. https://doi.org/10.4085/462-20.

56. Otto C, Barthel D, Klasen F, Nolte S, Rose M, Meyrose AK, et al. Predictors of self-reported health-related quality of life according to the EQ-5D-Y in chronically ill children and adolescents with asthma, diabetes, and juvenile arthritis: longitudinal results. Qual Life Res. 2018;27(4):879–90. https://doi.org/10.1007/s11136-017-1753-8.

57. Hsu CN, Lin HW, Pickard AS, Tain YL. EQ-5D-Y for the assessment of health-related quality of life among Taiwanese youth with mild-to-moderate chronic kidney disease. Int J Qual Health Care. 2018;30(4):298–305. https://doi.org/10.1093/intqhc/mzy011.

58. Nugawela MD, Stephenson T, Shafran R, De Stavola BL, Ladhani SN, Simmons R, et al. Predictive model for long COVID in children 3 months after a SARS-CoV-2 PCR test. BMC Med. 2022;20(1):465. https://doi.org/10.1186/s12916-022-02664-y.

59. Zhou W, Shen A, Yang Z, Wang P, Wu B, Herdman M, et al. Patient-caregiver agreement and test-retest reliability of the EQ-5D-Y-3L and EQ-5D-Y-5L in paediatric patients with haematological malignancies. Eur J Health Econ. 2021;22(7):1103–13. https://doi.org/10.1007/s10198-021-01309-w.

60. Amien R, Scott D, Verstraete J. The validity and reliability of the interviewer-administered EQ-5D-Y-3L version in young children. Health Qual Life Outcomes. 2023;21(1):19. https://doi.org/10.1186/s12955-023-02100-6.

61. Control CfD. Prevention. Measuring healthy days: Population assessment of health-related quality of life. Atlanta, Georgia; 2001. Available from: https://stacks.cdc.gov/view/cdc/6406/cdc_6406_DS1.pdf.

62. Arbuckle R, Abetz-Webb L. "Not just little adults": qualitative methods to support the development of pediatric patient-reported outcomes. The Patient. 2013;6(3):143–59. https://doi.org/10.1007/s40271-013-0022-3.

63. Amien R, Scott D, Verstraete J. Performance of the EQ-5D-Y Interviewer Administered Version in Young Children. Children. 2022;9(1):93. https://doi.org/10.3390/children9010093.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.