What Does the EQ-5D-Y-5L Measure in Comparison to other Generic Paediatric Health Related Quality of Life Instruments? A Dimensionality Assessment Using Factor Analysis

Authors: Mina Bahrampour[1], Renee Jones[2], Kim Dalziel[2], Nancy Devlin[3], Brendan Mulhern[1], on behalf of QUOKKA (Quality of Life in Kids: Key Evidence for Decision Makers in Australia) team

1-Centre for Health Economics Research and Evaluation (CHERE), University of Technology Sydney

2-Health Economics Unit, University of Melbourne

3-School of Population and Global Health, University of Melbourne

Abstract

Objectives:

The EQ-5D-Y-5L (Y-5L) includes the same five dimensions of health-related quality of life (HRQoL) as the adult version. The Y-5L is one of a number of widely used generic measures of child HRQoL, including the Health Utilities Index (HUI), Child health Utility (CHU9D) and Peadiatric Quality of Life Inventory (PedsQL). The content of each is different; yet there is little evidence on the measurement relationship between these instruments, and the additional domains of HRQoL added by each. This study aims to explore the measurement relationship between the Y-5L and other validated child HRQoL instruments, utilising unique data from the Australian Paediatric Multi-Instrument Comparison Study (P-MIC) [1].

Methods:

Data from the P-MIC study[2], were used. Y-5L, PedsQL, CHU9D and HUI 2/3 data were collected from both proxies and self-reported by children. Exploratory Factor Analysis (EFA) was used to investigate the underlying dimension structure. To assess overall dimensionality, the items from the four instruments were pooled, and modelled for child and proxy report separately. To build up a picture of the instrument relationships, the dimensionality between the Y-5L and each other instrument was examined. The suitability of data for EFA was checked using Kaiser-Meyer-Olkin (KMO) and Bartlett's test of sphericity. The number of factors was based on eigenvalues greater than one; eigenvalue measures the variance explained by each factor.

[1] Jones R, Mulhern B, McGregor K, Yip S, O'Loughlin R, Devlin N, Hiscock H, Dalziel K, On Behalf Of The Quality Of Life In Kids Key Evidence To Strengthen Decisions In Australia Quokka Project Team. Psychometric Performance of HRQoL Measures: An Australian Paediatric Multi-Instrument Comparison Study Protocol (P-MIC). Children (Basel). 2021; 8(8):714. doi: 10.3390/children8080714.

[2] Data cut 1, dated 6th May 2022.

The correlation coefficient for the variable and factor is shown by loadings in the analysis; this study used loadings more than 0.32, with cross-loading also considered. The factors were assumed to be correlated; hence oblique rotation (Promax) was used.

Results:

When applying EFA to the pooled model, additional domains not measured by the Y-5L can be identified. Results suggested a six-factor structure for the proxy data and a different but overlapping six-factor structure for the child self-report data. Factors related to mental health, pain and daily activities were covered by the Y-5L; however, additional factors related to social functioning, school functioning and senses (vision and hearing) were not.

Assessing instrument pairs suggests that PedsQL is sensitive to social functioning and school functioning issues not explicitly measured by Y-5L. The sensitivity of HUI to hearing and vision reflects the inclusion of these as separate items in it.

Conclusion:

We provide evidence about what the Y-5L measures in comparison to other instruments. This study provides information which can be used to guide choice of generic instrument when measuring particular domains of HRQoL. The results also suggest factors that may represent relevant candidate items for Y-5L bolt-ons. These both align with and differ from bolt-ons identified for the adult EQ-5D. The results support earlier research suggesting that factor analysis can be a useful statistical method for identifying potential gaps and new dimensions for EuroQol instruments.

Introduction:

Health related quality of life (HRQoL) is an important outcome indicator in the evaluation of healthcare interventions and treatments, and in allocating health resources (1). Generic measures of HRQOL, accompanied by preference weights, are commonly used in economic evaluation to measure HRQoL and inform the estimation of quality adjusted life years (QALYs) (2). QALYs measure health outcomes by combining the quality of life with the length of life. These instruments have two components, a descriptive (classification) system and a value set. The descriptive system usually comprises a number of dimensions and levels associated with them. The value set is a set of weights based on stated preferences assigned to each health state and used to provide the quality of life input of the QALY. It is anchored on a scale from one for full health to zero which is equal to being dead (3).

A range of preference-accompanied instruments have been developed for use in the adult population including the EQ-5D-3L (4), EQ-5D-5L (5), SF-6D v1(6, 7) and v2 (8). Studies have found that these instruments might not be suitable to be used in child and adolescent populations. Children have different understanding about health than adults, evidence suggests children focus on well-being and psychosocial health whereas adults(proxies) concentrate more on the absence of illness or disability (9). In other words, the domains of HRQoL important to a child and the most appropriate way to describe them might be different to that of an adult (10). Overall the adult instruments might not suitable to be used in the younger population in terms of dimensions, wording, response scale and formatting (11, 12); hence, descriptive systems of relevance to the younger population have been developed (12-14).

The EQ-5D-Y-3L (15) was developed to assess HRQoL in an age-appropriate manner, based on the same dimension structure as the adult EQ-5D instruments. Self-care is referred to as looking after self and anxiety/depression is worded as worried, sad and unhappy (16). The EQ-5D-Y-5L is an experimental five level version of the instrument (17). The EQ-5D-Y-3L and EQ-5D-Y-5L are part of a broader suite of paediatric specific measures of HRQoL (18). Other widely used and validated generic measures of child HRQoL include Health Utilities Index (HUI) (19), Child health Utility (CHU9D) (12) and Peadiatric Quality of Life Inventory (PedsQL) (20). Even though these instruments are all developed to measure generic HRQoL, they differ in important ways. They differ in terms of the HRQoL domains measured, or where they do

measure the same construct, they use different item wording or different response levels (statement versus likert). Second, they use different outcome scales (frequency versus severity) and different recall periods (one month versus today versus usual)[3].

To validate HRQoL instruments purporting to measure the same construct, it is possible to investigate what domains of the construct are being measured by each, and how the instruments converge and diverge in the overall assessment of HRQoL. If HRQoL is conceptualized as the overall construct that is being measured, then different approaches can be used to understand the measurement relationship between instruments within the overall construct. For example, five domains of HRQoL are measured by the EQ-5D-Y, but it is not known to what extent the domains of HRQoL measured by other instruments converge or diverge with those measured by the EQ-5D-Y. Understanding this allows for the further validation of instrument content, construct validity and understanding the measurement relationship across instruments.

Similar work has been conducted by Finch et al (21, 22), who found that there are some aspects of health that are not covered by the EQ-5D adult version but are included in other generic instruments, such as hearing. This evidence was used to support the identification of potential bolt-on dimensions.

However, work assessing dimensionality (the domains of HRQoL measured by a set of items from different instruments) across instruments has not been conducted for measures of paediatric HRQoL. Therefore, this study aims to explore the measurement relationship between four instruments for both proxy and self-complete using factor analysis. It builds on the work of Finch et al (21, 22) and provides evidence about what the EQ-5D-Y-5L measures in comparison with other paediatric instruments. We also identify specific dimensions missing from the EQ-5D-Y-5L compared with those measured by other common paediatric HRQoL instruments.

Methods:

[3] There are currently no value sets available to accompany either the EQ-5D-Y-5L or the PedsQL, although methodological studies preparatory to providing such values are underway for both. PedsQL does have published utility mapping functions to the EQ-5D-Y-3L and CHU-9D.

Data source: Data from the paediatric multi- instrument (P-MIC) study data cut 1, dated 6th May 2022, were used (23). This study is part of the wider Quality of Life in Kids: Key Evidence for Decision Makers in Australia (QUOKKA) research programme in Australia. A detailed summary of the P-MIC data collection is available from the study protocol (19), and the paper presented by Jones et al (24) at this plenary meeting. The P-MIC study collects a number of different generic and condition-specific instruments, a subset of which are included this study. Four child-specific generic HRQoL measures (EQ-5D-Y-5L, PedsQL, CHU9D and HUI2/3) were included in the comparison. These instruments were included as they were under consideration for valuation as part of the QUOKKA programme. As valuation of the EQ-5D-Y-3L is already underway in Australia, it was not included in the decision-making process, and was therefore not included in this study. The EQ-5D-Y-5L, PedsQL and CHU9D were administered to the whole sample as these instruments as part of the core instrument set, while HUI2/3 was administered to a subset (approximately one third) of the online panel to minimize respondent burden (25). The data collection for the P-MIC is still ongoing; however, the vast majority of data has been collected by the time this study has been done (data cut 1, May 6th 2022).

Instruments measured:

EQ-5D-Y-5L: The EQ-5D-Y-5L measures five dimensions of health using single items, able to walk around, looking after myself, doing usual activities, having pain or discomfort, and feeling worried, sad, or unhappy. Each dimension assesses severity across five severity response levels ranging from no problems to unable to/extreme problems (5).

***CHU-9D*:** The CHU9D has nine dimensions and each dimension has 5 severity response categories (from 'no' (don't feel) to 'very' in five items; and 'no problem' to 'can't do' in four items). The domains are worry, sadness, pain, tiredness, annoyed, school, sleep, daily routine and joining in with activities (12, 14).

***HUI*:** The HUI Mark 2 and 3 (HUI2/3) instrument can be used to classify an individual's HRQoL according to wither the HUI 2 classification system or HUI 3 classification system. As the HUI 3 classification system was developed to address issues in the HUI 2 system, the HUI 3 classification system was used. The HUI 3 classification system has eight dimensions measured across 15 items, each with five or six response levels. The domains are vision (2

items), hearing (2 items), speech (2 items), ambulation (2 items), dexterity (1 item), emotion (2 items), cognition (2 items) and pain (2 items) (26, 27).

***PedsQL*:** The PedsQL Generic Core 4.0 was used which is composed of 23 items that measure four broad dimensions defined as physical, emotional, social and school functioning. Each item has five frequency levels (from never, to almost always) (28).

Data analysis

Convergence assessment using correlations: To provide a basis for the factor analysis and dimensionality assessment, the convergence and divergence between the EQ-5D-Y-5L and each of the other instrument items was assessed using correlations. Spearman correlation was used as the distribution was not normal. If correlation scores were less than 0.3 the correlation was considered weak, if scores were between 0.3 and less than 0.5 the correlation was moderate, and scores of 0.5 or higher indicated a strong correlation (29).

Factor Analysis:

Choice of factor analysis approach*:* There are different methods to detect the dimensionality in variables such as Principal Components Analysis (PCA) and Exploratory Factor Analysis (EFA). Even though both methods can be used to reduce the number of variables, the main aim of each method is slightly different; PCA is a technique for reducing the dimensionality of the data, whereas EFA is a method for identifying and measuring latent variables or factors, which cannot be measured directly. As we wanted to investigate the dimensionality of the items and identify the latent factors, EFA was used to investigate the underlying dimensional structure. To assess the overall dimensionality, the items from the four instruments were pooled for the self-complete and proxy report data separately. The dimensionality between the EQ-5D-Y-5L and each other instrument was also assessed. This was done to understand the measurement relationship between each instrument. It also allowed for comparisons with the overall pooled model by allowing for an assessment of how the addition of further instruments changed the dimension pattern.

Data check for suitability*:* Kaiser-Meyer-Olkin (KMO) and Bartlett's test of sphericity were used to assess the suitability of the data for EFA. KMO tests examine the strength of the partial correlation between items. A rule of thumb for interpreting the KMO is, that values between

0.8 and 1 indicate the sample is adequate to run a factor analysis. The significance level for Bartlett's test should be below 0.05. A p-value less than 0.05 on Bartlett's test indicates that individual variables are sufficiently correlated for a factor analysis to be accomplished.

Choosing factors and items: When using EFA, some components may help to decide the best structure for the item presented. The method can be decided based on the number of factors included in the model, items representing each factor, and the correlation between Items.

The number of factors in EFA can be decided upon by the number of eigenvalues. The eigenvalue shows the variance which the factors explain. The rule of thumb for choosing factors is based on eigenvalues greater than one, but factor structures can be forced to extract a certain number of factors. This can be supported by the scree plots (that plot the eigenvalues) and parallel analysis (determines the number of factors based on eigenvalues). To check all the possible factor structures, 6, 7, 8, 9, 10, and 11 factors were tested in this study.

Each factor consists of items; to choose the best items, loadings are used. Loadings are the correlation between item and factor and uniqueness is the variance that is unique to that item in the model and it is equal to (1- loading^2). Factors and domains have been used interchangeably in this study. Loadings more than 0.32 were chosen to keep the items (30, 31) for each factor (using the same rule as the correlation mentioned above). If there was cross-loading, i.e., if an item had a loading greater than 0.32 on two factors, we chose the factor with the higher loading. As we assumed there might be a correlation between factors, oblique rotation (Promax) was used. All the analyses were done using Stata software version 16.0 (32).

Results:

Sample: Of the 6247 participants available in the P-MIC data cut, 5444 were children aged 5 to 18 years of age and hence were include in this analysis. Of the 5444, a total of 1039 children and adolescents fully completed four instruments (EQ-5D-Y-5L, PedsQL, CHU9D, and HUI), of which 548 were parent/proxy report. For the three instruments completed by everyone (EQ-5D-Y-5L, PedsQL, CHU9D), there were 3525 fully completed by children and adolescents of which 1919 were parent/proxy report. The child age for the proxy completed instruments was

between 5 to 18 years and the age of the child for the self-complete was 7 to 18 years children and adolescents.

The data were tested for suitability prior to running the EFA. The KMO was 0.928 and 0.951 for proxies and self-complete data respectively and both groups had a significant Bartlett test (p-value <0.001).

Table 1 Demographic summary of respondents completing all four instruments

	Proxy complete (n=548)	**Self-complete(n=1039)**
	frequency (percent)	**frequency (percent)**
Age		
5-6	264(48.14%)	-
7-10	112 (20.44%)	440 (42.34%)
11-15	117 (21.35%)	394 (37.92%)
16-18	55 (10.03%)	205 (19.73%)
Gender		
Male	292 (52.28%)	553 (53.22%)
Female	249 (45.44%)	473 (45.52%)
other	7 (1.09%)	13 (1.25%)
Need or use medicine prescribed by a doctor		
No	390 (71.17%)	712 (68.53%)
Yes	158 (28.83%)	327 (31.47%)

Convergent validity

Correlation between EQ-5D-Y-5L, and CHU-9D, PEDSQL and HUI: Table 2 reports the correlations (with the highest coefficients for each item that is greater than 0.5 displayed in **bold**).

CHU9D: The correlation pattern between EQ-5D-Y-5L and CHU9D in both proxy and self-report groups was almost identical and the same items had the highest correlation, except sleeping domain in CHU9D, which had the highest correlation with EQ-5D-Y-5L domain feeling worried sad or unhappy for the self-complete whereas in the proxy group sleeping had the highest correlation with usual activities. In both groups, pain had the highest correlation between dimensions that could be hypothesised to correlate. Most of the CHU9D items correlated closely with the EQ-5D-Y-5L feeling worried, sad or unhappy.

PedsQL: The results indicated a low to moderate correlation between the EQ-5D-Y-5L and *PedsQL* items. The correlation pattern was different between the two groups. The highest correlation in the self-completed group was between "sad" from PedsQL and "feeling worried, sad or unhappy" from EQ-5D-Y-5L. The highest correlation in the proxy completed surveys was between "taking a bath" from PedsQL and "looking after themselves" from EQ-5D-Y-5L.

HUI: The correlation between the instruments shows a low to moderate correlation. Different correlation patterns can be seen between the two groups. The items related to pain had the highest correlation in the self-completed surveys whereas in the proxy completed the highest correlation was between the "perform basic activities" and "looking after themselves".

Table 2 Correlation between EQ-5D-Y-5L and the three other instruments (strong correlation, more than 0.5, bolded)

	Self-complete					**Proxy complete**				
EQ-5D-Y-5L	**MO**	**SC**	**UA**	**PD**	**AD**	**MO**	**SC**	**UA**	**PD**	**AD**
CHU-9D										
Worried	0.227	0.239	0.380	0.314	**0.667**	0.168	0.234	0.370	0.314	**0.639**
Sad	0.236	0.258	0.409	0.322	**0.610**	0.164	0.224	0.360	0.280	**0.599**
Pain	0.396	0.223	0.373	**0.696**	0.327	0.349	0.215	0.330	**0.739**	0.299
Tired	0.223	0.200	0.344	0.356	0.452	0.204	0.210	0.321	0.299	0.380
Annoyed	0.206	0.256	0.359	0.297	0.499	0.128	0.272	0.364	0.222	0.465
School work	0.206	0.311	0.423	0.276	0.421	0.233	0.419	**0.526**	0.258	0.385
Sleeping	0.239	0.256	0.398	0.351	0.471	0.187	0.309	0.396	0.293	0.393
Daily routine	0.312	**0.544**	0.519	0.299	0.398	0.319	**0.618**	0.605	0.286	0.397
Join Activities	0.312	0.358	**0.576**	0.337	0.435	0.321	0.498	**0.639**	0.344	0.407
PedsQL										
Walk	0.461	0.362	0.414	0.367	0.289	0.364	0.365	0.356	0.239	0.180
Run	0.412	0.282	0.396	0.397	0.321	0.383	0.328	0.346	0.303	0.220
Sport activities	0.415	0.324	0.463	0.389	0.348	0.354	0.380	0.434	0.293	0.253
Lifting	0.300	0.321	0.343	0.300	0.269	0.310	0.346	0.343	0.250	0.175
Taking bath	0.352	**0.617**	0.435	0.237	0.273	0.258	**0.588**	0.429	0.176	0.178
Doing chores	0.290	0.441	0.468	0.299	0.374	0.219	0.441	0.420	0.188	0.237
Hurting	0.351	0.282	0.397	**0.627**	0.357	0.273	0.253	0.322	**0.514**	0.310
Low energy	0.307	0.246	0.415	0.435	0.456	0.270	0.224	0.333	0.339	0.354
Scared	0.192	0.287	0.343	0.321	**0.558**	0.098	0.287	0.317	0.227	0.496
Sad	0.202	0.247	0.389	0.347	**0.630**	0.140	0.219	0.362	0.292	**0.585**
Angry	0.148	0.279	0.343	0.267	0.476	0.115	0.295	0.362	0.194	0.420
Sleep	0.226	0.259	0.381	0.374	0.478	0.187	0.304	0.397	0.282	0.398
Worry	0.188	0.227	0.351	0.364	**0.604**	0.129	0.190	0.306	0.265	**0.559**
Getting along	0.189	0.304	0.375	0.251	0.423	0.104	0.230	0.281	0.140	0.221

Other kids playing	0.184	0.282	0.354	0.223	0.403	0.161	0.331	0.378	0.194	0.326
Getting teased	0.172	0.279	0.319	0.225	0.403	0.139	0.242	0.318	0.237	0.349
Not able to keep-up	0.322	0.394	0.463	0.314	0.389	0.353	0.526	**0.561**	0.299	0.335
Keeping up	0.330	0.375	0.474	0.325	0.406	0.260	0.392	0.408	0.240	0.217
Paying attention	0.205	0.317	0.422	0.289	0.431	0.113	0.305	0.321	0.131	0.173
Forgetting	0.207	0.291	0.390	0.296	0.387	0.189	0.360	0.428	0.239	0.336
Schoolwork	0.211	0.316	0.420	0.291	0.417	0.135	0.306	0.364	0.171	0.229
Missing school unwell	0.247	0.230	0.401	0.388	0.376	0.242	0.240	0.378	0.347	0.295
Missing school hospital	0.297	0.288	0.399	0.369	0.320	0.305	0.341	0.428	0.357	0.286
HUI										
Vision 1	0.052	0.056	0.108	0.115	0.139	0.077	0.118	0.152	0.091	0.184
Vision 2	0.118	0.131	0.124	0.114	0.139	0.064	0.131	0.183	0.060	0.132
Hearing 1	0.247	0.181	0.156	0.211	0.125	0.154	0.230	0.189	0.251	0.122
Hearing 2	0.231	0.125	0.153	0.126	0.119	0.266	0.174	0.174	0.270	0.110
Speech 1	0.224	0.291	0.329	0.212	0.266	0.219	0.404	0.419	0.218	0.249
Speech 2	0.239	0.262	0.327	0.196	0.229	0.277	0.435	0.408	0.134	0.208
Feel 1	0.223	0.210	0.377	0.286	**0.545**	0.202	0.284	0.437	0.290	0.486
Pain 1	0.323	0.206	0.338	**0.633**	0.384	0.343	0.243	0.324	**0.623**	0.366
Walk	**0.520**	0.308	0.309	0.262	0.186	0.499	0.343	0.311	0.323	0.248
Use hands	0.333	0.269	0.165	0.170	0.079	0.379	0.318	0.285	0.152	0.157
Remember	0.222	0.246	0.355	0.233	0.359	0.138	0.333	0.409	0.191	0.349
Think	0.253	0.334	0.470	0.299	0.442	0.227	0.485	**0.556**	0.245	0.423
Perform basic activities	0.372	**0.510**	0.359	0.225	0.224	0.322	**0.677**	0.550	0.222	0.270
Feel 2	0.241	0.240	0.388	0.386	**0.641**	0.143	0.310	0.430	0.321	**0.589**
Pain 2	0.292	0.215	0.280	**0.593**	0.352	0.270	0.177	0.285	**0.642**	0.346

Dimensionality Assessment

Pooled item model – Self Report: The best fitting model includes six factors with eigenvalues greater than one. Two of PedsQL items, "doing Chores" and "missing school hospital" did not load on any factor. The variance explained is 88%. Figure 1 presents the domain structure identified, with the factor loadings for each item included. The domains are defined as:

1. *Emotional Functioning* which includes one EQ-5D-Y-5L item worried, sad or unhappy, alongside seven PedsQL, six CHU-9D and two HUI items. The factor loadings range from 0.331 to 0.822. The items with the highest loadings are EQ-5D-Y-5L "worried, sad

[4] some of the item names have been added as abbreviation, which the full definition can be found in the appendix

and unhappy", and "worried", and "sad" from the CHU-9D. The items with the lowest loadings are focused on the emotional impact of relationship with other children.

2. *Daily activities* which includes three items from the EQ-5D-Y-5L, five HUI items, and one items each from PedsQL and CHU9D. The item with the highest loading is EQ-5D-Y-5L "looking after self", followed by the PedsQL item focused on washing. The usual activities items from the other instruments are also included alongside broader activities related constructs such as dexterity and communication.
3. *Cognition/ school functioning* does not include any of the EQ-5D-Y-5L items, it includes the "school work" item from CHU9D, five items from PedsQL and two items from HUI 3. Factor loading range from 0.352 to 0.911, the lowest loading is related to missing school item from PedsQL.
4. *Pain* this domain includes all the items related to pain, which includes one item from EQ-5D-Y-5L , one item from CHU9D, item "hurting" from PedsQL and two HUI items. The loading range from 0.490 to 0.642. The item with the lowest loading is the PedsQL "hurting".
5. *Physical functioning* all items loading on this domain are from PedsQL. Item ranges from 0.407 and 0.799. The items with the lowest loadings are focused on the keeping up with other children.
6. *Senses* includes items related to hearing and vision, higher loadings were related to "vision" and had a loading greater than 0.7 whereas the loading for "hearing" items were less than 0.4.

Of note here is that the EQ-5D-Y-5L is included in three of the six dimensions (emotional functioning, daily activities and pain), and is not included in physical functioning (demonstrating that the EQ-5D-Y-5L "walking around" item does not have a strong relationship with the PedsQL physical functioning items), cognition and school functioning, and senses. Four items had cross-loadings, PedsQL "taking a bath" with physical functioning domain, "angry" with cognition and school functioning domain, "other kids playing" and "hurting" with physical functioning domain.

The results indicated that modelling seven or more factors divides items related to the same domain such as emotional functioning or senses into multiple domains, and the items related to missing school appear as a separate domain.

***Pooled item model – Proxy report*:** The best fitting model also includes six factors with eigenvalues greater than one. All PedsQL items except "missing school hospital" loaded for this model. The variance explained is 86%. Figure 2 presents the domain structure identified, with the factor loadings for each item included. The domains are defined as:

1. *Emotional Functioning* which includes one EQ-5D-Y-5L item "worried, sad or unhappy", alongside six PedsQL, five CHU-9D and two HUI items. The factor loadings range from 0.465 to 0.810. The item with the highest loading is EQ-5D-Y-5L "worried, sad or unhappy".
2. *Daily activities* which includes three items from the EQ-5D-Y-5L, alongside three items from CHU9D, 1 item from PedsQL and 7 items from HUI. The loading ranges from 0.359 which is PedsQL "not able to keep up" item , to 0.836 which is "looking after themselves" form EQ-5D-Y-5L.
3. *Cognition/ school functioning* includes seven items from PedsQL, the highest loading is 0.878 for "school activities" and the lowest loading is 0.400 for "getting teased".
4. *Pain* this domain includes all the items related to pain, which includes one item from EQ-5D-Y-5L , one item from CHU9D, item "hurting" from PedsQL and two HUI items. The loading range from 0.490 to 0.642. The item with the lowest loading is the PedsQL "hurting". In the proxy report EFA there was an additional item in pain domain "Missing school feeling unwell".
5. *Physical functioning* items were all from PedsQL. This domain did not involve the relation between children and their peers like self report. The loading range was from 0.388 to 0.856, the item with the lowest loading was "doing chores".
6. *Senses* includes items related to hearing and vision, loading range was from 0.489 to 0.731, with item related to "vision" having a higher loadings.

Similar to the self-report results, in the proxy report results the EQ-5D-Y-5L is included in three of the six dimensions (emotional functioning, daily activities and pain), and is not included in physical functioning (demonstrating that EQ-5D-Y-5L. "walking around" from EQ-5D-Y-5L had a cross loading on pain domain (loading 0.347). Three other items that had cross loadings, "taking a bath" with daily activities domain, "keeping up" with physical functioning domain, and "missing school unwell" with pain domain.

Figure 1. Conception model of EFA results for items pooled from all instruments (self- complete) and item loadings

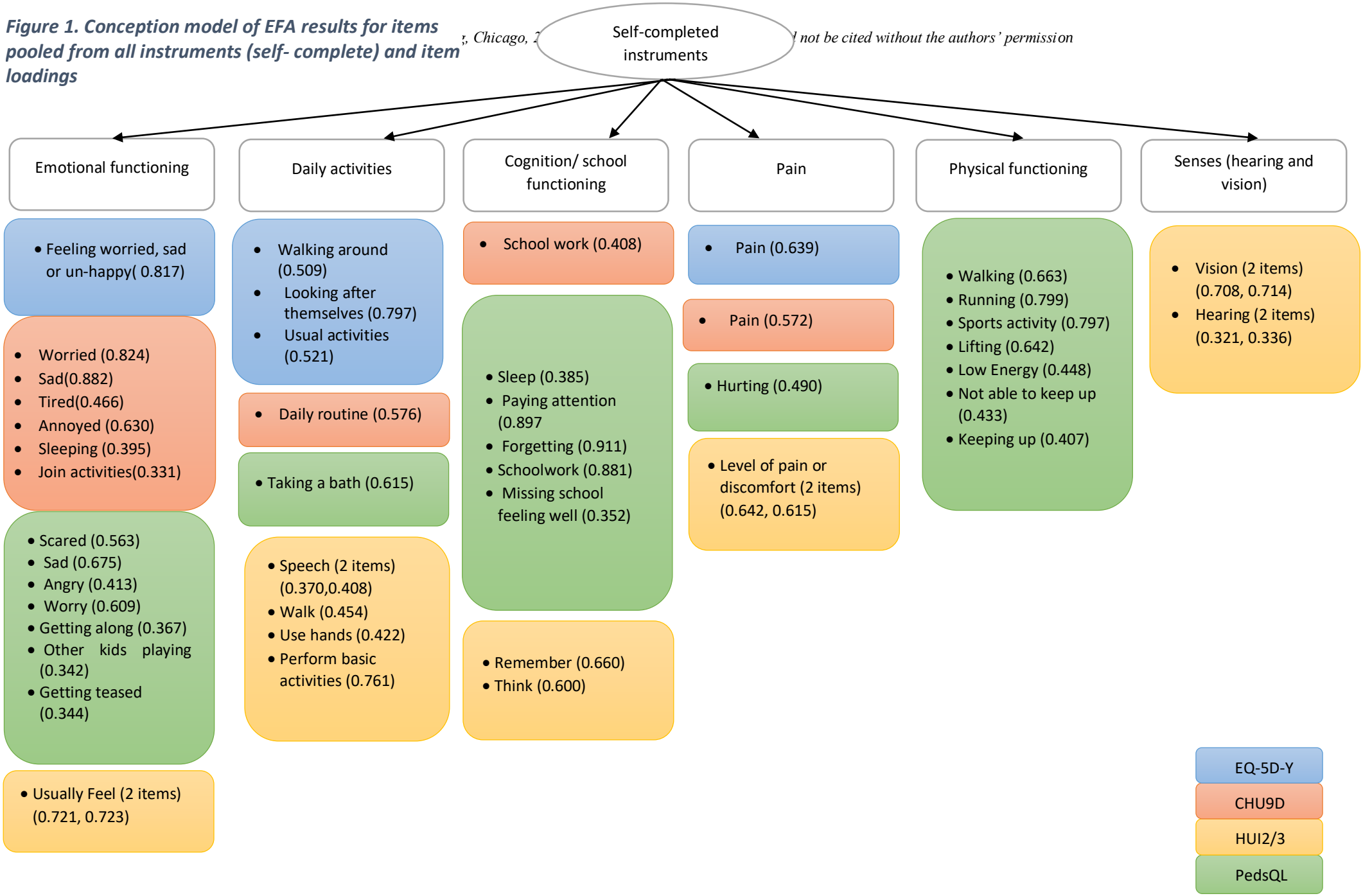
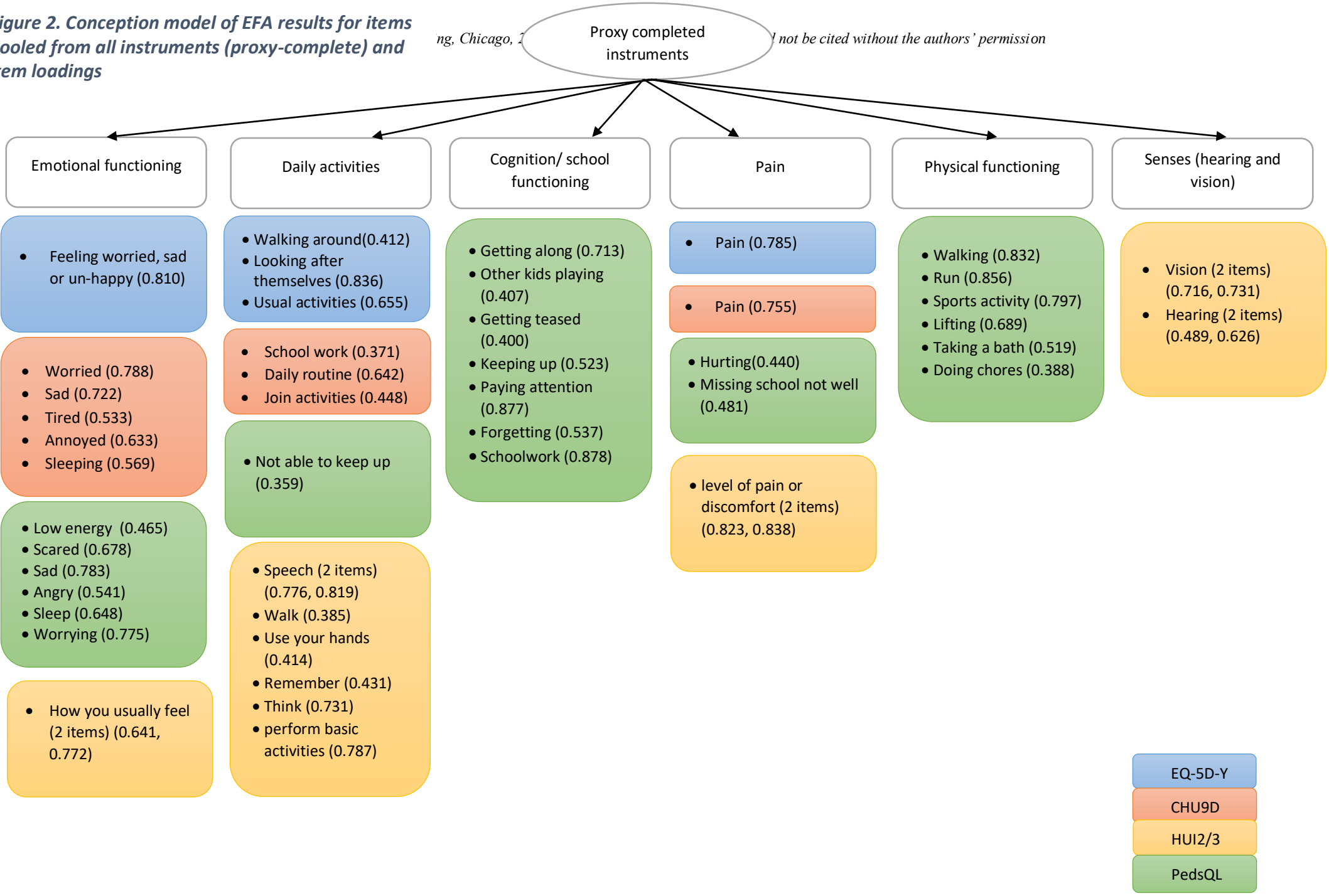


Figure 2. Conception model of EFA results for items pooled from all instruments (proxy-complete) and item loadings



Dimensionality of EQ-5D-Y-5L and CHU9D: The results of EFA between EQ-5D-Y-5L and CHU9D, showed in the self-complete group all the items loaded on one factor, which explained 78% of the model variation (Table 3, where the blanks are loadings less than 0.3). And in the proxy group, items loaded on two factors, which together explained 96% of the model variation. Pain from CHU9D loaded on both factors, but its loading was higher on factor two which involves more items related to emotional functioning.

Table 3 EFA results between EQ-5D-Y-5L and CHU9D

		Self		Proxy		
Instruments	**Variable**	**Factor 1**	**Uniqueness**	**Factor 1**	**Factor 2**	**Uniqueness**
EQ-5D-Y-5L	Walking around	0.507	0.743	0.672		0.653
	Looking after themselves	0.532	0.717	0.941		0.326
	Usual activities	0.703	0.505	0.818		0.287
	Pain	0.647	0.582	0.441		0.616
	Feeling worried, sad or unhappy	0.741	0.452		0.814	0.356
CHU9D	Worried	0.694	0.518		0.838	0.385
	Sad	0.718	0.485		0.868	0.355
	Pain	0.635	0.597	0.303	0.336	0.673
	Tired	0.626	0.609		0.491	0.673
	Annoyed	0.631	0.602		0.637	0.568
	School work	0.571	0.674	0.506		0.573
	Sleeping	0.648	0.580		0.418	0.636
	Daily routine	0.670	0.552	0.747		0.375
	Join activities	0.648	0.580	0.605		0.472

Dimensionality of EQ-5D-Y-5L and PedsQL*:* The EFA result from the self-completed instruments showed a three-factor structure in which 90% of the variation is explained by these factors. The proxy reports show a four-factor structure with 95% of the model explained by these factors. EQ items loaded across two domains even given the additional explanatory power of the PedsQL. The proxy-report model adds two factors that do not include any EQ-5D-Y items, which are items from school functioning and physical functioning. (Table 4)

Dimensionality of EQ-5D-Y-5L and HUI*:* The EFA result from the self and proxy reports both showed a four-factor structure adding a senses domain to EQ-5D-Y-5L, which includes items assessing hearing and vision. The variance explained by the self-report was 96% whereas this was 85% in the proxy report. (Table 5)

Table 4 EFA results for EQ-5D-Y-5L and PedsQL

		Self				Proxy				
Instruments	**Variable**	**Factor1**	**Factor2**	**Factor3**	**Uniqueness**	**Factor1**	**Factor2**	**Factor3**	**Factor4**	**Uniqueness**
EQ-5D-Y-5L	Walking around	0.720			0.579			0.670		0.561
	Looking after themselves	0.523	0.307		0.599			0.867		0.327
	Usual activities	0.504			0.552			0.819		0.319
	Pain	0.536		0.394	0.529			0.526		0.606
	Feeling worried, sad or un-happy			0.614	0.494	0.725				0.481
PedsQL	Walk	0.825			0.344		0.874			0.325
	Run	0.808			0.344		0.961			0.225
	Sport activities	0.799			0.298		0.835			0.257
	Lifting	0.584			0.576		0.721			0.433
	Taking bath	0.568	0.329		0.526		0.421			0.520
	Doing chores	0.405	0.334		0.534		0.381			0.574
	Hurting	0.540		0.420	0.441	0.425	0.313			0.572
	Low energy	0.401		0.480	0.432	0.510	0.329			0.538
	Scared			0.647	0.434	0.739				0.499
	Sad			0.764	0.294	0.868				0.336
	Angry		0.310	0.548	0.488	0.550				0.597
	Sleep			0.567	0.514	0.531				0.597
	Worry			0.742	0.386	0.873				0.388
	Getting along		0.713		0.394				0.672	0.536
	Other kids playing		0.778		0.384	0.348			0.412	0.534
	Getting teased		0.719		0.444	0.446			0.328	0.578
	Not able to keep-up	0.377	0.559		0.373			0.454		0.403
	Keeping up	0.367	0.566		0.356				0.529	0.467
	Paying attention		0.639		0.404				0.887	0.328
	Forgetting		0.551		0.483				0.476	0.484
	Schoolwork		0.636		0.414				0.841	0.313
	Missing school unwell			0.369	0.568	0.385				0.592
	Missing school hospital	0.347			0.591			0.355		0.569

Table 5 EFA results for EQ-5D-Y-5L and HUI2/3

		Self					Proxy				
Instruments	**Variable**	**Factor1**	**Factor2**	**Factor3**	**Factor4**	**Uniqueness**	**Factor1**	**Factor2**	**Factor3**	**Factor4**	**Uniqueness**
EQ-5D-Y-5L	Walking around	0.567				0.552	0.432	0.439			0.539
	Looking after themselves	0.669				0.526	0.767				0.345
	Usual activities	0.426	0.338			0.531	0.595				0.375
	Pain			0.553		0.490		0.728			0.346
	Feeling worried, sad or un-happy		0.644			0.457			0.548		0.456
HUI2/3	Vision 1				0.766	0.434				0.810	0.419
	Vision 2				0.761	0.432				0.821	0.409
	Hearing 1					0.829				0.484	0.702
	Hearing 2					0.821				0.565	0.527
	Speech 1	0.483	0.322			0.523	0.792				0.432
	Speech 2	0.488				0.562	0.832				0.388
	Feel 1		0.687			0.460			0.645		0.463
	Pain 1			0.723		0.337		0.843			0.334
	Walk	0.565		0.319		0.571	0.420	0.337			0.538
	Use hands	0.501				0.721	0.449				0.733
	Remember		0.549			0.635	0.379		0.466		0.574
	Think		0.636			0.434	0.660		0.433		0.320
	perform basic activities	0.729				0.492	0.774				0.464
	Feel 2		0.728			0.358			0.755		0.306
	Pain 2			0.711		0.368		0.838			0.328

Discussion:

The study reports the use of EFA to understand the measurement relationship across four commonly used paediatric HRQoL instruments for both caregiver/proxy and child self-report. This study builds on earlier work assessing the dimensionality of item pools, and is unique in applying EFA to explore the domain structure of HRQoL instruments used for paediatric populations and to compare the results between self and proxy reports.

The same number of domains resulted from the proxy and self-completed instruments, however, the domains that were identified differed. Items related to senses (vision and hearing) was similar in both groups; items related to cognition from HUI loaded on daily activities domain in proxy completed, but they loaded on cognition and school activities domain for the self-completed instruments, this shows that proxies may relate "doing usual activities" to include cognition, but children consider their cognitive function as something else while answering item related to the cognitive ability (33). Also items related to socializing with peers from PedsQL such as "other kids not wanting to play with him or her" and "getting teased by other children" loaded on emotional functioning for self-report whereas in the proxy-completed data these items loaded on cognition and school activities domain. This shows that children and proxies may have different understandings regarding HRQoL items, which is in line with previous research (34, 35).

The results have implications for the assessment of HRQoL using the EQ-5D-Y-5L. For the pooled model, items from EQ-5D-Y-5L load onto three domains: emotional functioning, pain, and daily activities. Other domains resulting from EFA were physical functioning, cognitive/school functioning, and senses. The new domains show how other instruments broaden what is measured by the EQ-5D-Y-5L. These added domains can be useful especially where the EQ-5D-Y-5L might not be sensitive enough to capture HRQoL related to some conditions or populations. The results also indicate that EQ-5D-Y-5L does not measure five distinct domains, but combines the physical items together, with separate coverage of mental health, and in certain models, pain, and this has also been indicated elsewhere for the adult instrument (36). Similarly, the CHU9D, which was developed as a nine dimension preference based measure mainly loads onto the emotional functioning. This may indicate that the CHU9D is more appropriate for assessing emotional impacts of HRQoL. However, when interpreting the results it is important to take into account the differences in item wording and response levels.

Another possible issue might be how items interrelate and affect different aspects of health even though they load on the same domains (36-38).

Cognitive functioning is related to HRQoL in children and adolescents (33). The majority of frequently used generic HRQoL instruments for children have a dimension related to cognition. Even though EQ-5D-Y-5L was developed for this population, it does not have a cognition dimension. Cognitive functioning, which is mostly related to concentration and school functioning, was a domain resulting from the EFA in both proxy and self-completed results.

Another factor that has resulted from the EFA is senses which contains items regarding hearing and vision. The effect of adding hearing and vision has also been investigated in the adult EQ-5D (39): adding bolt-on items for vision and hearing had a significant impact on values of the health states of the EQ-5D adult version, although the direction and magnitude of differences depended on the severity of the health state (40). The effect of adding such item to the youth version can be an avenue for further research.

The results from this study have important implications for the development and validation of potential bolt-on dimensions for the EQ-5D-Y. The domains that were not covered by any of the EQ-5D-Y-5L items are cognition function (33), school activities (which are loaded on the same factor), and senses. The identification of factors and items is the first important step for deriving bolt-ons systematically. This study can be a step to further research in developing bolt-ons for the EQ-5D youth versions. Ludwig et al stated that including a cognitive dimension in the EQ-5D-Y-3L improves the measurement of HRQoL (33). The study also indicated that adding a cognitive dimension to EQ-5D-Y will give it a similar structure as PedsQL; the PedsQL has been widely used to estimate HRQoL in the paediatric population and is widely used and accepted by paediatricians[5] (28). Cognition functioning has been examined and recommended as an important dimension to be added as a bolt-on to the adult EQ-5D (41), and adding this dimension showed a significant impact on health state values for this instrument.

The limitation of this study is that we did not use Confirmatory Factor Analysis (CFA). This is because the aim of the study was not to reduce items to find the best items to develop a new instrument but to find new domains and better understand what other instruments add to

[5] As noted earlier, PedsQL is not currently accompanied by preference weights, but methodological work to address that is currently underway

what is measured by the EQ-5D-Y-5L, therefore we did not delete any items that had a loading greater than 0.32. Also, the results of this study can be used as a base theory for CFA in bigger data sets to find the best structures to be used in both groups, adding new domains as a bolt-on can help with more accurate results, especially in different disease/condition groups. Other limitations of this study are that it does not include disease-specific measures in the analysis, this might cause not including some domains which are important for some conditions. Also the focus of this paper was on EQ-5D-Y-5L and not EQ-5D-Y-3L, however the research team will run the analysis to compare the EQ-5D-Y-3L with other three instruments to see if there are any differences in the EFA results.

Conclusion:

This study provides new evidence regarding HRQoL domains covered by different commonly used paediatric generic instruments, and how these instruments measure similar or different things. For example, cognition and school functioning is captured by all instruments in the study except EQ-5D-Y. The same domains resulted from proxy and self-reports, however items presenting the domains differ which shows the different views of children and the proxies. Results suggest that EFA can be a useful statistical method to identify add-on domains to HRQoL instruments. This method, in addition to other quantitative and qualitative methods can be applied to help identify potential gaps and new dimensions for potential bolt-ons for EQ-5D-Y especially for children who have a certain disease or condition.

Appendix:

Appendix A : Abbreviation table

Items full name	Abbreviation used in paper
EQ-5D-Y-5L	Item | Domain
Walking around	*Walking around | MO*
Looking after themselves	*Looking after themselves |SC*
Usual activities	*Usual activities| UA*
Pain	*Pain| PD*
Feeling worried, sad or un-happy	*Feeling worried, sad or un-happy |AD*
CHU-9D	
Worried	*Worried*
Sad	*Sad*
Pain	*Pain*
Tired	*Tired*
Annoyed	*Annoyed*
School work	*School work*
Sleeping	*Sleeping*
Daily routine	*Daily routine*
Join Activities	*Join Activities*
PedsQL	
Walking more than one block	*Walk*
Running	*Run*
Participating in sports activity or exercise	*Sport activities*
Lifting something heavy	*Lifting*
Taking a bath or shower by him or herself	*Taking bath*
Doing chores around the house	*Doing chores*
Having hurts or aches	*Hurting*
Low energy level	*Low energy*
Feeling afraid or scared	*Scared*
Feeling sad or blue	*Sad*
Feeling angry	*Angry*
Trouble sleeping	*Sleep*
Worrying about what will happen to him or her	*Worry*
Getting along with other children	*Getting along*
Other kids not wanting to be his or her friend	*Other kids playing*
Getting teased by other children	*Getting teased*
Not able to do things that other children his or her age can do	*Not able to keep-up*
Keeping up when playing with other children	*Keeping up*
Paying attention in class	*Paying attention*
Forgetting things	*Forgetting*
Keeping up with schoolwork	*Schoolwork*
Missing school because of not feeling well	*Missing school unwell*
Missing school to go to the doctor or hospital	*Missing school hospital*
HUI	
Ability to see well enough to read ordinary newsprint	*Vision 1*

Ability to see well enough to recognize a friend on the other side of the street	*Vision 2*
Ability to hear what was said in a group conversation with at least three other people,	*Hearing 1*
Ability to hear what was said in a conversation with one other person in a quiet room	*Hearing 2*
Ability to be understood when speaking your own language with people who do not know you	*Speech 1*
Ability to be understood when speaking with people who know you well	*Speech 2*
How you usually feel	*Feel 1*
Level of pain or discomfort	*Pain 1*
Ability to Walk	*Walk*
Ability to use your hands and fingers	*Use hands*
Ability to remember things	*Remember*
Ability to think and solve day to day problems	*Think*
Ability to perform basic activities	*perform basic activities*
How you usually feel	*Feel 2*
Level of pain or discomfort	*Pain 2*

References:

1. Kromm SK, Bethell J, Kraglund F, Edwards SA, Laporte A, Coyte PC, et al. Characteristics and quality of pediatric cost-utility analyses. Quality of Life Research. 2012;21(8):1315-25.
2. Richardson J, Schlander M. Health technology assessment (HTA) and economic evaluation: efficiency or fairness first. J Mark Access Health Policy. 2018;7(1):1557981-.
3. Prieto L, Sacristán JA. Problems and solutions in calculating quality-adjusted life years (QALYs). Health and Quality of Life Outcomes. 2003;1(1):80.
4. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Annals of medicine. 2001;33(5):337-43.
5. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727-36.
6. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. Journal of health economics. 2002;21(2):271-92.
7. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. Med Care. 2004;42(9):851-9.
8. Brazier JE, Mulhern BJ, Bjorner JB, Gandek B, Rowen D, Alonso J, et al. Developing a New Version of the SF-6D Health State Classification System From the SF-36v2: SF-6Dv2. Medical Care. 2020;58(6):557-65.
9. Ungar W. Economic evaluation in child health: Oxford University Press; 2009.
10. Petrou S. Methodological issues raised by preference-based approaches to measuring the health status of children. Health economics. 2003;12(8):697-702.
11. Eiser C, Morse R. Quality-of-life measures in chronic diseases of childhood. Health technology assessment (Winchester, England). 2001;5(4):1-157.
12. Stevens KJ. Working with children to develop dimensions for a preference-based, generic, pediatric, health-related quality-of-life measure. Qualitative health research. 2010;20(3):340-51.

13. Dalziel K, Catchpool M, García-Lorenzo B, Gorostiza I, Norman R, Rivero-Arias O. Feasibility, Validity and Differences in Adolescent and Adult EQ-5D-Y Health State Valuation in Australia and Spain: An Application of Best–Worst Scaling. PharmacoEconomics. 2020;38(5):499-513.
14. Stevens K, Ratcliffe J. Measuring and Valuing Health Benefits for Economic Evaluation in Adolescence: An Assessment of the Practicality and Validity of the Child Health Utility 9D in the Australian Adolescent Population. Value in Health. 2012;15(8):1092-9.
15. 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.
16. 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. Quality of life research. 2010;19(6):887-97.
17. Kreimeier S, Åström M, Burström K, Egmar A-C, Gusi N, Herdman M, et al. EQ-5D-Y-5L: developing a revised EQ-5D-Y with increased response categories. Quality of Life Research. 2019;28(7):1951-61.
18. Kwon J, Freijser L, Huynh E, Howell M, Chen G, Khan K, et al. Systematic Review of Conceptual, Age, Measurement and Valuation Considerations for Generic Multidimensional Childhood Patient-Reported Outcome Measures. PharmacoEconomics. 2022;40(4):379-431.
19. Horsman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI®): concepts, measurement properties and applications. Health and Quality of Life Outcomes. 2003;1(1):54.
20. Varni JW, Burwinkle TM, Seid M. The PedsQL™ as a pediatric patient-reported outcome: Reliability and validity of the PedsQL™ Measurement Model in 25,000 children. Expert review of pharmacoeconomics & outcomes research. 2005;5(6):705-19.
21. Finch AP, Brazier JE, Mukuria C, Bjorner JB. An Exploratory Study on Using Principal-Component Analysis and Confirmatory Factor Analysis to Identify Bolt-On Dimensions: The EQ-5D Case Study. Value in Health. 2017;20(10):1362-75.
22. Finch AP, Brazier JE, Mukuria C. Selecting Bolt-On Dimensions for the EQ-5D: Examining Their Contribution to Health-Related Quality of Life. Value in Health. 2019;22(1):50-61.
23. Jones R, Mulhern B, Devlin N, Hiscock H, O'Loughlin R, McGregor K, et al. Australian Paediatric Multi-Instrument Comparison (P-MIC) Study: Technical Methods Paper [Online]. 2022.
24. Renee Jones BM, Nancy Devlin, Rachel O'Loughlin, Xiuqin Xiong, Mina Bahrampour, Kristy McGregor, Shilana Yip, Harriet Hiscock, Kim Dalziel on behalf of QUOKKA (Quality of Life in Kids: Key Evidence for Decision Makers in Australia). The Australian paediatric multi-instrument comparison (P-MIC) study: data quality, feasibility, acceptability, and descriptive comparisons of the EQ-5D-Y-3L, EQ-5D-Y-5L, TANDI and PedsQL. EuroQol Group Plenary; Chicago 2022.
25. 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).
26. Gemke RJ, Bonsel GJ. Reliability and validity of a comprehensive health status measure in a heterogeneous population of children admitted to intensive care. Journal of clinical epidemiology. 1996;49(3):327-33.
27. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, et al. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Medical care. 2002;40(2):113-28.
28. Varni JW, Seid M, Kurtin PS. PedsQL™ 4.0: Reliability and validity of the Pediatric Quality of Life Inventory™ Version 4.0 Generic Core Scales in healthy and patient populations. Medical care. 2001:800-12.
29. Cohen J. A power primer. Psychological bulletin. 1992;112(1):155-9.
30. Tabachnick BG, Fidell LS, Ullman JB. Using multivariate statistics: pearson Boston, MA; 2007.
31. Hair JF, Black WC, Babin BJ, Anderson RE, Tatham RL. Multivariate data analysis (Vol. 6). 2006.

32. StataCorp. 2019. Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC.
33. Ludwig K, Surmann B, Räcker E, Greiner W. Developing and testing a cognitive bolt-on for the EQ-5D-Y (Youth). Quality of Life Research. 2022;31(1):215-29.
34. Jardine J, Glinianaia SV, McConachie H, Embleton ND, Rankin J. Self-reported quality of life of young children with conditions from early infancy: a systematic review. Pediatrics. 2014;134(4):e1129-e48.
35. Pickard AS, Topfer L-A, Feeny DH. A structured review of studies on health-related quality of life and economic evaluation in pediatric acute lymphoblastic leukemia. JNCI Monographs. 2004;2004(33):102-25.
36. Gamst-Klaussen T, Gudex C, Olsen JA. Exploring the causal and effect nature of EQ-5D dimensions: an application of confirmatory tetrad analysis and confirmatory factor analysis. Health and quality of life outcomes. 2018;16(1):153-.
37. Feng YS, Jiang R, Kohlmann T, Pickard AS. Exploring the Internal Structure of the EQ-5D Using Non-Preference-Based Methods. Value Health. 2019;22(5):527-36.
38. Shah KK, Mulhern B, Longworth L, Janssen MF. Views of the UK General Public on Important Aspects of Health Not Captured by EQ-5D. The patient. 2017;10(6):701-9.
39. Yang Y, Rowen D, Brazier J, Tsuchiya A, Young T, Longworth L. An exploratory study to test the impact on three "bolt-on" items to the EQ-5D. Value Health. 2015;18(1):52-60.
40. Longworth L, Yang Y, Young T, Mulhern B, Hernández Alava M, Mukuria C, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. Health technology assessment (Winchester, England). 2014;18(9).
41. Krabbe PFM, Stouthard MEA, Essink-Bot M-L, Bonsel GJ. The Effect of Adding a Cognitive Dimension to the EuroQol Multiattribute Health-Status Classification System. Journal of Clinical Epidemiology. 1999;52(4):293-301.