This is a repository copy of *Selection and validation of a classification system for a child-centred caries-specific utility measure*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/160762/

Version: Published Version

Monograph:

Rogers, H.J., Gilchrist, F., Marshman, Z. et al. (2 more authors) (2020) Selection and validation of a classification system for a child-centred caries-specific utility measure. Report. ScHARR HEDS Discussion Papers . School of Health and Related Research, University of Sheffield

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

The University Of Sheffield.

HEALTH ECONOMICS & DECISION SCIENCE

Discussion Paper Series

Title: Selection and validation of a classification system for a child-centred caries-specific utility measure

Authors: Rogers HJ, Gilchrist F, Marshman Z, Rodd HD, Rowen D

Corresponding author: Helen J Rogers, Unit of Oral Health, Dentistry and Society, School of Clinical Dentistry, University of Sheffield, UK.
Email: hrogers1@sheffield.ac.uk

Selection and validation of a classification system for a child-centred caries-specific utility measure

Rogers HJ,[1] Gilchrist F,[1] Marshman Z,[1] Rodd HD,[1] Rowen D[2]

1) Unit of Oral Health, Dentistry and Society, School of Clinical Dentistry, University of Sheffield, UK

2) Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, UK

Corresponding author: Helen J Rogers (ORCID: 0000-0001-5031-5009), Unit of Oral Health, Dentistry and Society, School of Clinical Dentistry, University of Sheffield, UK. Email: hrogers1@sheffield.ac.uk

Abstract

Background: Caries Impacts and Experiences Questionnaire for Children (CARIES-QC) is a child-centred caries-specific quality of life measure. This study aimed to select, and validate with children, a classification system for a paediatric condition-specific preference-based measure, based on CARIES-QC.

Methods: First, a provisional classification system for a preference-based measure based on CARIES-QC was developed using Rasch analysis, psychometric testing, involvement of children and parents, and the developer of CARIES-QC. Second, qualitative, semi-structured 'think aloud' validation interviews were undertaken with a purposive sample of children with dental caries. The interviewer aimed to identify whether items were considered important and easily understood, whether any were overlapping and if any excluded items should be reintroduced. Interview recordings were transcribed verbatim and thematic analysis conducted.

Results: Rasch analysis identified poor item spread for the items *'cross'* and *'school'*. Items relating to eating were correlated and the better performing items were considered for selection. Children expressed some confusion regarding the items *'school'* and *'food stuck'*. Parent representatives thought that impacts surrounding toothbrushing (*'brushing'*) were encompassed by the item *'hurt'*. Five items were selected from CARIES-QC for inclusion in the provisional classification system; *'hurt'*, *'annoy'*, *'carefully'*, *'kept awake'* and *'cried'*. Validation interviews were conducted with 20 children aged 5-16 years old. Participants thought the questionnaire was straightforward and covered a range of impacts. Children thought an item about certain foods being *'hard to eat'* was more relevant than one about having to eat more carefully because of their teeth and so the *'carefully'* item was replaced with *'hard to eat'*.

Conclusion: Following child-centred modification, the preliminary five-item classification system is considered valid and suitable for use in a valuation survey. The innovative child-centred methods used to both identify and validate the classification system can be applied in the development of other preference-based measures.

Key words: caries, children, oral health-related quality of life, utility

<u>**Declarations**</u>

Background

Dental caries (tooth decay) is a prevalent oral disease, causing significant negative impacts on the lives of children and young people [1,2]. Whilst pain is the most common feature of caries, there is a growing body of evidence on the further impacts relating to pain on children's daily lives [3]. These include time off school, difficulty sleeping, speaking, eating and interference with everyday activities [4,5]. Furthermore, a number of studies have highlighted links between dental caries and general health, with higher levels of untreated dental caries reported to be associated with reduced weight and poor growth [6,7].

The wider impacts of caries on society are also substantial. In the United Kingdom, approximately 57,485 children aged up to 19-years were admitted to hospital in 2015-2016 with a diagnosis of dental caries [8]. As a result, dental caries remains the most common reason for children to require hospital admission with an estimated cost of £39 million to the National Health Service (NHS) [8].

Dental caries is a largely preventable disease and a range of community-based programmes and clinical strategies have been adopted to reduce the prevalence in children. However, there have been few high quality economic evaluations to determine the cost effectiveness of such programmes [9-11]. This creates difficulties for decision-makers and commissioners in determining which interventions to provide within the remit of the National Health Service [12,13].

Within child oral health research, this paucity of economic evaluations could be attributed to the lack of a suitable instrument to measure Quality Adjusted Life Years (QALYs) [14]. Presently, only one generic preference-based measure has been used in child oral health research, with limited success; the Child Health Utility 9D (CHU9D) was not found to be sensitive enough to changes in caries status [14]. The lack of use of other measures and the poor psychometric performance of CHU9D suggests that the content of child and adolescent generic preference-based measures may not be appropriate or sensitive for use in oral health research.

There is a clear need for the development of a validated preference-based measure, specifically for children, that is appropriate for measuring treatment benefits for dental caries [14]. This is achievable through the adaptation of a novel child-centred caries-specific oral health-related quality of life (OHRQoL) measure, known as CARIES-QC (Caries Impacts and Experiences Questionnaire for Children) [15]. This measure was developed with involvement of children at every stage, addressing the primary limitation of a number of other measures of OHRQoL [16]. In its current form, CARIES-QC cannot be used to generate QALYs since it is not preference-

based. A preference-based measure consists of: a) a classification system that is used to describe the health of all children; and b) a value set used to score all health states described by the classification system.

It is not feasible or practical to gain preference weights from children for all of the twelve items within CARIES-QC [17]. As such, it was necessary to reduce the number of items within the measure to identify the classification system. Furthermore, the preliminary classification system would warrant validation with children prior to use in a valuation survey.

This study aimed to identify a classification system for a child-centred preference-based measure using a combination of statistical methodologies and involvement of stakeholders including children, young people and parents. Furthermore, this study sought to validate, and refine where necessary, the preliminary classification with a sample of children who had experience of dental caries using a qualitative approach. The valuation of the classification system to generate a preference-based measure will be reported elsewhere.

Methods

Ethical approval for this study was obtained from Yorkshire and the Humber Research Ethics Committee (Reference: 18/YH/0148).

Identification of the classification system

Several condition-specific preference-based measures have been developed using a staged approach that selects the classification system using a combination of Rasch Analysis, classical psychometric analysis and developer input ([18-21]). The present study builds upon this approach by also incorporating child and parent views. The following stages were used to identify the most appropriate items for a classification system:

1. Rasch Analysis
2. Classical psychometric analysis
3. Patient and Public Involvement (PPI)
4. Developer input

The study team discussed the findings of each approach, particularly where stakeholder views were found to conflict with the results of statistical analyses. Where this occurred, agreement was sought by consensus on which items should be selected for inclusion in the preliminary classification system. The final part of this study involved the validation of the preliminary classification system using a qualitative approach. Revisions to the classification system were undertaken accordingly.

CARIES-QC

The CARIES-QC is a 12-item measure (Table 1) that seeks children's assessment of the severity of their caries-related impacts, and has been validated for use in 5-16 year-olds. The response format of this measure differs from other measures of OHRQoL in that the three levels ('not at all', 'a bit' and 'a lot') relate to the severity, rather than the frequency, of impacts [16]. A response of 'a bit' would be assigned one point, 'a lot' would score two points, whilst 'not at all' suggests the impact has not been experienced and hence a no points are assigned. A total score can then be calculated. This instrument has been shown to have good face, content and construct validity, responsiveness and reliability [15]. Furthermore, the involvement of children at every stage during the development of CARIES-QC addresses an acknowledged need to view children as active participants within research [22].

Table 1: The questions within CARIES-QC (excluding the global question), the related items and severity levels

Questions from CARIES-QC	Items		Levels
How much do your teeth hurt you?	Hurt		Not at all
Do your teeth make it hard to eat some foods?	Hard to eat		
Do you have to eat on one side of your mouth because of your teeth?	Eating on one side		
Do you get food stuck in your teeth?	Food stuck		
How much do you get kept awake by your teeth?	Kept awake		A bit
How much do your teeth annoy you?	Feeling annoyed		
How much do your teeth hurt when you brush them?	Hurt when brushing		
Do you have to eat more carefully because of your teeth?	Eat more carefully		
Do you have to eat more slowly because of your teeth?	Eat more slowly		A lot
Do you feel cross because of your teeth?	Feeling cross		
How much have you cried because of your teeth?	Cried		
Do your teeth make it hard to do your schoolwork?	Difficulty doing schoolwork		

Data set

The data set for this study came from the original validation study for the CARIES-QC measure, which has been published elsewhere [15]. The data were from a sample of 200 children aged 5 to 16 years who had a diagnosis of active dental caries. Children were asked to complete the CARIES-QC measure at three different timepoints: baseline (T0), prior to the start of treatment (T1) and following a course of dental treatment to manage the caries (T2). Whilst all timepoints were used in the original validation of CARIES-QC, the present study used data from

timepoint T0 on which to conduct psychometric and Rasch analyses, as this had the highest number of observations with no attrition. A range of clinical data were also collected to establish the number of teeth affected by caries, whether children reported symptoms from their teeth, and the pattern of caries (i.e. whether it affected the front teeth) [15].

1. Rasch Analysis

Rasch Analysis has been used to convert each participant response onto a latent continuous scale representing the severity of impacts relating to OHRQoL and assesses the spread of responses across the three response levels for each item [18]. Items with a greater spread indicate that the responder is able to distinguish between the item levels and would be stronger candidates for inclusion in the classification system.

In this study, the Rasch Analysis focussed on the spread of items across the three levels (response categories) at logit 0, whereby a greater spread indicated the respondent was able to distinguish between the item levels. Item (x^2) goodness-of-fit statistics were also conducted, with the items having the best fit to the underlying model being the best candidates for inclusion in the classification system. Item fit residual scores were also applied in the same way, with those closest to 0 indicating a better fit to the model.

Differential Item Functioning (DIF) was also assessed to determine whether each item was working the same across respondents of different ages, genders, ethnicities and levels of deprivation according to Index of Multiple Deprivation (IMD) scores [23]. Threshold analyses and assessment of local dependencies were also conducted.

Rasch Analysis was conducted using RUMM2030™ software Version 5.3 (©Rumm Laboratory Pty Ltd., Perth, Australia).

2. Classical Psychometric Testing

Classical psychometric analyses were carried out using SPSS® software (IBM Corporation., New York, United States, Version 24)[18]. Initial Factor analysis was carried out alongside subgroup analyses relating to the age, gender and ethnicity of participants. This was followed by analyses to determine the rate of missing data, floor and ceiling effects, as well as correlations between items and with clinical findings, such as the number of decayed (D), missing (M) and filled (F) teeth (T) (known collectively as the DMFT index)[24-27].

Items that were more appropriate for selection for inclusion in the classification system were those with low missing data, as well as items with low floor effects (since the item would not be able to capture a deterioration in health). Items with ceiling effects were considered for selection as these can indicate that they capture the

impacts of higher disease severity. Correlations were used to identify items that were capturing the same aspect of quality of life, where one of the items may be selected in the classification system to reflect the wider set of items.

3. Views of Patient and Public Involvement (PPI) representatives

A panel of children and young people comprising personal contacts and local schoolchildren and patients from a paediatric dental clinic were invited to give their views at one of two informal meetings held in May and July 2017, to determine their views on the items within CARIES-QC. The panel was comprised of children from a range of ages, genders and ethnicities, with differing experiences of dental caries. This panel were also involved as a steering group for the overall study. These discussions focussed on how important each item was felt to be, whether any items were considered to overlap, and whether any items were felt to be too similar. Two parent representatives were also involved in these discussions, to provide their thoughts on the items within CARIES-QC from their perspectives.

4. CARIES-QC development insights

The fourth stage of this process centred on informal discussions with researchers involved in the development of CARIES-QC. It was important to acknowledge any issues or concerns identified by the research team during the development of this instrument, particularly since children were involved at every stage. Furthermore, it was essential that any difficulties surrounding the use of the instrument in different settings and languages were considered.

The findings from these four steps informed the generation of the preliminary classification system (version 1).

Child-centred validation of the preliminary classification system

Validation of the preliminary classification system was undertaken with children and young people who had a diagnosis of dental caries. Potential participants were identified via referral letters received from general dental practitioners at the Paediatric Dental Clinic at the Charles Clifford Dental Hospital, Sheffield. Patients with known diagnoses of dental caries were approached following their initial examination at the dental hospital. A maximum variation purposive sampling approach was used, to ensure participants of different ages, genders, ethnicities and socio-economic status. Participants were not eligible for inclusion if they were outside of the 5- to 16-year-old age range within which CARIES-QC has been validated. Furthermore, children and parents who were unable to understand spoken and written English language were excluded. A similar approach was used in both formulating

the descriptive system and testing the content validity of CARIES-QC [28]. Based on this previous research, it was expected that approximately 20 interviews would be required to reach data saturation.

Parents and children were invited to consent and assent to participate respectively. Qualitative semi-structured interviews were conducted by an experienced qualitative researcher (HJR). A topic guide (see Supplementary material) was used to inform the interviews, which were recorded and transcribed verbatim. Children were asked to 'think aloud' whilst completing questions from CARIES-QC within the preliminary classification system whilst the interviewer aimed to determine whether items were considered important, easily understood, and whether any were overlapping [29]. Children were then shown items that were excluded from the preliminary classification system and questioned further to determine whether any should be reintroduced.

Further sociodemographic data, including participant age and ethnicity were also collected. Postcodes were documented to facilitate calculation of the Index of Multiple Deprivation for each participant, given the well-acknowledged relationship between caries experience and socioeconomic deprivation [30,23]. Clinical caries experience was recorded for each participant using the aforementioned DMFT index [31].

Simple descriptive statistics were undertaken on the quantitative data. Qualitative data were analysed using the framework method to inform validation of the classification system, using NVivo 12 (©QSR International Pty Ltd) software for data management. This latter analysis focussed on identifying childrens' level of understanding for each item, the amount of importance participants placed upon each item and whether they considered any as redundant or overlapping. PPI representatives for the study were involved in confirming the interpretation of quotes from children and young people were correct. The study team discussed the qualitative findings, which were used to inform modification of the preliminary classification system as required.

Results

Identification of the classification system

1. Rasch Analysis

The 200 participants from the aforementioned CARIES-QC validation dataset were included in the Rasch model. The sociodemographic characteristics and caries experience of the participants in this dataset are provided in Table 2. Eleven participants did not fit the Rasch model and so were removed from further Rasch analysis. Overall, the CARIES-QC data were found to have a good item (mean 0.385 ± 0.902) and person fit (Mean 0.254 ± 0.999) to the Rasch model.

Regarding the individual items, none were found to have disordered thresholds (Figure 1), and none were subjected to local dependency (less than 0.2 above the average correlation)[32]. The thresholds differed for the items *'food stuck'* and *'school'*, as seen in Figure 1.

Table 2: Sociodemographic characteristics and caries experience of participants from the original CARIES-QC validation study (dataset used to undertake Rasch analysis and classical psychometric testing in the present study) and the qualitative validation of the preliminary classification system derived from CARIES-QC

	CARIES-QC validation dataset used for Rasch and classical psychometric analyses (*n*=200)		Qualitative validation of preliminary classification system (*n*=20)	
Age (years)	Mean: 8.1	Range: 5-16	Mean: 10.1	Range: 6-15
Gender				
Male	95	(47.5%)	6	(30.0%)
Female	105	(52.5%)	14	(70.0%)
Ethnicity				
Asian background	31	(15.5%)	2	(10.0%)
Black background	5	(2.5%)	1	(5.0%)
Mixed background	9	(4.5%)	2	(10.0%)
White British background	130	(65.0%)	14	(70.0%)
Other background	9	(4.5%)	1	(5.0%)
Unknown background	16	(8.0%)	0	(0.0%)
Socioeconomic status				
Most deprived	119	(59.5%)	10	(50.0%)
More deprived	37	(18.5%)	0	(0.0%)
Average	20	(10.0%)	3	(15.0%)
Less deprived	13	(6.5%)	3	(15.0%)
Least deprived	11	(5.5%)	4	(20.0%)
Total dmft	Mean: 6.24 (SD: 3.45)	Range: 0-16	Mean: 2.85 (SD: 3.05)	Range: 0-12
Total DMFT	Mean: 1.57 (SD: 2.18)	Range: 0-13	Mean: 1.7 (SD: 2.88)	Range: 0-11

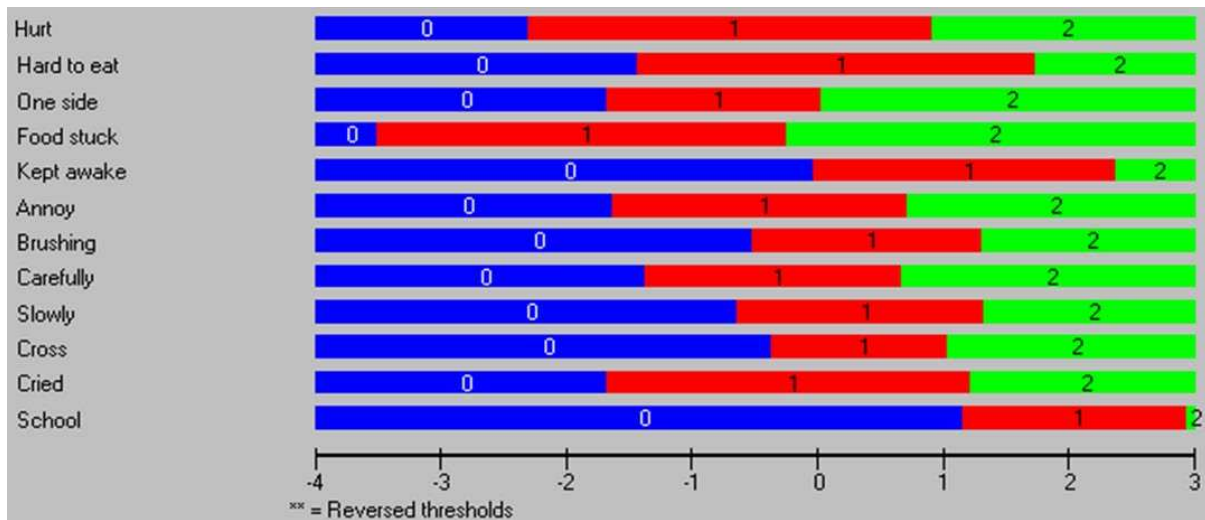*Figure 1: Threshold map for the items within CARIES-QC.*

Table 3 reports the results of the Rasch analysis. The item spread at logit 0 of the items 'cross' and 'school' were poor at 0.705 and 0.894 respectively.

The items '*annoy*' and '*carefully*' were found to have high negative item fit residuals (-1.802 and -1.801 respectively) and the item '*cried*' was found to have a high positive fit residual (1.112). Whilst these are notable, and could potentially indicate item redundancy (associated with Item-Total Correlation), a level of +/-2.5 should normally be reached for this to cause concern. '*Annoy*' was also found to not fit the Rasch model at the 5% level (p= 0.013).

The items '*hard to eat*' (0.031) and '*cross*' (0.021) were found to have uniform differential item functioning (DIF) with regard to age at the 5% level. '*Hard to eat*' also showed non-uniform DIF (0.014) at this level, as did '*eat on one side*' (0.049).

The item '*food stuck*' did not fit the model at the 5% significance level (p=0.036) and appeared to be working differently for variations in age groups (F=-0.293) and genders (F=-0.126).

2. Classical Psychometric Testing

Classical psychometric tests were undertaken on the same dataset used for the Rasch analysis (Table 2). There were moderate levels of correlation (between 0.3 and 0.5) between most items within CARIES-QC (see Supplementary material) [26]. Strong correlations (between 0.5 and 0.9) were found between the item '*annoy*' and five other items, namely '*hurt*' (0.59)*, 'one side'* (0.58)*, 'kept awake'* (0.52)*, 'carefully'* (0.55)*,* and *'cross'* (0.51). Similarly the item *'carefully'* had strong correlations with four other items, namely '*hard to eat* (0.51)*',*

'one side' (0.63)*, 'annoy'* (0.55)*, and 'slowly'* (0.60)*.* This suggests that not all of these items are needed within the classification system, which should include a smaller number of items to reflect what is captured by the wider measure. Correlations between each item and clinical indicators, such as the presence of decayed, missing and filled teeth (DMFT) did not reveal any issues.

The item *'food stuck'* was the only item to have a floor effect (32% responded 'a lot'), which suggests it may have been misinterpreted by respondents, or that a high proportion of children experienced this impact regardless of whether they had caries or not. A high ceiling effect (>50%) was noted for '*kept awake'* and '*cross'*, with 67% and 59% of respondents reporting no experience of these impacts. This indicates that these items may represent the more severe impacts of caries, which are likely to be experienced by fewer participants. Similarly, the items '*brushing'* and *'slowly'* had high ceiling effects.

No items had missing values for each item greater than 2.5%, and there were no salient findings arising from subgroup analyses [27].

3. Views of Patient and Public Involvement (PPI) representatives

Children and young people noted that there were multiple items within CARIES-QC relating to eating, and many participants suggested that one item alone could encompass the others on this topic. Children thought the items *'carefully'* and *'hard to eat'* had the broadest remit, and that one of these could be considered in place of the rest.

Children expressed some uncertainty about whether the item *'food stuck'* related to getting food stuck in their teeth in general, or getting food stuck in the holes in their teeth.

Children felt the term *'annoy'* was too similar to 'cross'. Older children in particular thought they would be less likely to use the word *'cross'*, and hence would prefer the item *'annoy'*.

Older children thought that their peers would not be likely to admit to crying about their teeth.

Child and parent representatives expressed some confusion about how schoolwork could be affected by teeth. They reasoned that if dental pain was causing the impacts on schoolwork, this may be captured elsewhere under the category of *'hurt'*.

Parent representatives thought that pain related to toothbrushing, could also come under the umbrella term *'hurt'*. They also considered whether *'hurt'* and *'annoy'* might mean the same thing, though children and young people disagreed.

4. CARIES-QC development insights

The item 'food stuck' has had translatability concerns when translating into other languages. Anecdotal evidence suggests that children may have a varied understanding of the schoolwork item. These two items could be excluded from the classification on this basis.

Children and young people of different ages viewed the concepts of 'hurt' and 'annoy' to be different during development of CARIES-QC, although both terms were used to describe the physical sensations that they felt. This suggests it may be important to retain both of these items within the preliminary classification system. In the qualitative research undertaken during the development of the CARIES-QC older children had admitted to crying about their teeth, in contrast to the suggestion made by the PPI representatives.

Table 3: Summary of key results from Rasch Analysis, classical psychometric testing, involvement of PPI representatives and discussions with the developers of CARIES-QC

Item	Item Spread at Logit 0	Item level fit Chi Squared (P-value)	Differential Item Functioning (DIF)	Residual	Disordered thresholds	Missing Data (%)	Floor effects (%)	Ceiling effects (%)	Strong correlations with other items	Concerns from PPI reps	Concerns from CARIES-QC development
Hurt	1.605	5.142 (0.076)	✗	-0.757	✗	✓ (1.5)	✓ (17% 'a lot')	✓ (31% 'not at all')	✓ (annoy)	✗	✗
Hard to eat	1.585	1.288 (0.525)	✓ (age* and gender**)	0.048	✗	✓ (2)	✗	✓ (43% 'not at all')	✓ (carefully)	✗	✗
Eating on **one side**	0.858	0.868 (0.648)	✓ (age**)	-0.793	✗	✓ (2.5)	✓ (25% 'a lot')	✓ (37% 'not at all')	✓ (annoy; carefully)	✓	✗
Food stuck	1.632	6.646 **(0.036)**	✓ (ethnicity* **)	0.661	✗	✓ (2)	✓ (32% 'a lot')	✗	✗	✓	✓
Kept awake	1.202	0.612 (0.736)	✗	-0.393	✗	✓ (1.5)	✗	✓ (67% 'not at all')	✓ (annoy)	✗	✗

Item											
Feeling **annoy**ed	1.174	8.699 **(0.013)**	✗	-1.802	✗	✓ (2)	✓ (18% 'a lot')	✓ (40% 'not at all')	✓ (hurt; one side; kept awake; carefully; cross)	✗	✗
Hurt when **brushing**	0.913	1.362 (0.506)	✓ (ethnicity**)	0.379	✗	✓ (0.5)	✗	✓ (57% 'not at all')	✗	✓	✗
Eat more **carefully**	1.019	4.367 (0.113)	✗	-1.801	✗	✗	✓ (18% 'a lot')	✓ (43% 'not at all')	✓ (hard to eat; one side; annoy; slowly)	✗	✗
Eat more **slowly**	0.988	1.775 (0.412)	✓ (deprivation*)	-0.874	✗	✓ 0.5	✗	✓ (55% 'not at all')	✓ (carefully)	✓	✗
Feeling **cross**	0.705	2.368 (0.306)	✓ (age* and ethnicity**)	0.130	✗	✗	✗	✓ (59% 'not at all')	✓ (annoy)	✓	✗
Cried	1.466	4.237 (0.120)	✗	**1.112**	✗	✗	✗	✓ (38% 'not at all')	✗	✓	✗
Difficulty doing **school**work	0.894	1.339 (0.512)	✓ (ethnicity*)	-0.536	✗	✓ (0.5)	✗	✓ (82% 'not at all')	✗	✓	✓

*Uniform DIF ** Non-uniform DIF

Discussion of preliminary classification system

The findings from all four steps outlined above were discussed between all members of the study team, and the preliminary classification system was agreed by consensus. A summary of the key discussion points is provided below, based upon the results seen in Table 3.

The items 'food stuck' and 'school' had issues noted in each of the four steps detailed above, and hence were excluded from the preliminary classification system. As the PPI representatives expressed a need for only one item relating to eating within the classification system, it was felt that 'eat more carefully' would encompass this best. This was in part due to its strong correlations with other items regarding impacts and experiences from eating, and its relatively good fit with the Rasch model. Similarly, the item 'annoy' was considered important to retain, given its strong correlations with clinical findings. Although parents expressed concerns that 'annoy' could be too similar to 'hurt', these items appeared to be independent of each other when analysing the data, and in previous qualitative research children considered them to be separate concepts during the development of the measure [33]. The items 'cried' and 'kept awake' were considered to be key components of the preliminary classification system, in order to represent the worst states.

Table 4 shows the five items were selected to form the preliminary classification system, and the broad domains represented by each. The preliminary five-item classification system was then ready for validation with children and young people.

Validation of the preliminary classification system

'Think aloud' interviews were conducted with 20 participants, of which 6 were male, and 14 female, before data saturation was reached. Two potential participants declined to take part; one parent felt their child was too shy to take part, whilst the other reported a lack of time.

The sociodemographic characteristics and caries experience of participants in shown in Table 2. The majority of participants (n=14) were White British, whilst the rest (n=6) identified with a variety of different ethnicities. The age of participants ranged from 6 to 15 years with a mean of 10 years. Using home postcodes to gain Index of Multiple Deprivation scores, half of the participants (n=10) were found to reside in the 'most deprived' areas of England. All children had active dental caries. The clinical caries experience of the sample was determined by calculating the number of decayed, missing and filled teeth in the primary dentition (dmft) and permanent dentition

(DMFT) for each participant. The mean dmft was 2.85 (SD 3.05; range 0-12) and DMFT was 1.7 (SD 2.88; range 0-11). The mean length of interview was 8 minutes and 10 seconds, though this ranged from below 5 minutes to upwards of 16 minutes, with the shortest interviews involving younger children.

The qualitative findings arising from the validation of the classification system are described below, with quotes provided to illustrate each aspect, using participant pseudonyms

Complexity

Children found the questions relating to the preliminary classification system straightforward to complete and did not appear to experience much difficulty in choosing an answer for each question. Furthermore, they believed the questions covered a range of impacts.

"They're kind of easy…but they mean a lot" Jenny, 11 years old.

Children were unsure whether their school friends would be able to answer some of the questions that had been removed from the classification system.

On questioning, younger children struggled to make decisions between items and found it difficult to communicate a clear preference for items capturing similar aspects of health:

"*Both…I like them both*" Lucy, 6 years old.

Overlapping items

During the development of the preliminary classification system, parent representatives for the study had raised some concern that the items 'hurt' and 'annoy' were too similar and potentially overlapping. Nonetheless, these interviews suggest the contrary, as children felt 'hurt' and 'annoy' described different things, and considered them both to have value.

"I think they're very different because annoying and hurt are two different meanings" Ali, 13 years old.

Importance of items

Children had conflicting views on the item 'cried' relating to the question 'have you ever cried because of your teeth'. Those who had experienced this impact placed greater importance on this item:

"'Cause sometimes if they really hurt, I do cry…..I actually think that is important" Lucy, 6 years old.

However, those who had never experienced this impact expressed confusion:

"I don't really know why people would cry about their teeth" Lily, 14 years old.

Appropriateness of items

Children thought the question 'do you have to eat more carefully because of your teeth?' did not adequately describe the dietary restrictions resulting from caries. They displayed a clear preference for one of the questions that had been removed from the classification system, which asked whether their teeth made it hard to eat some foods.

"*If you eat more carefully you can still eat but if you find it hard to eat you can't really eat much*" Leon, 9 years old.

"Because if you have to eat more carefully it's like how you eat whereas "Does your teeth make it hard to eat some foods?" would like eliminate foods out." Lily, 14 years old.

Child-centred modification of preliminary classification system

The findings from the qualitative interviews were then used to inform modifications to the preliminary classification system accordingly.

During the validation interviews, children raised some important issues with the item regarding 'eating more carefully', particularly that it failed to encompass their dietary limitations due to caries. They expressed a clear preference for the item 'hard to eat', and thought this item should be reinserted in the place of the problematic item. The rest of the items within the preliminary classification system were easily understood and considered to be both important and appropriate. Furthermore, children believed the items to be independent of each other, and not overlapping. The final validated classification system can be seen in Table 4.

Table 4: The preliminary classification system and final validated classification system, with proposed domains

Preliminary classification system	Final validated classification system	Proposed domain
Hurt	Hurt	Physical impacts
Annoy	Annoy	Physical impacts
Carefully	Hard to eat	Impacts on daily activities
Kept awake	Kept awake	Impacts on sleep

| Cried | Cried | Emotional impacts |

Discussion

This paper describes a novel approach to identify a classification system for a paediatric condition-specific preference-based measure from a condition-specific patient-reported outcome measure. The approach taken here builds on the previous approach taken to select items for many condition-specific preference-based measures through the validation of the classification system using qualitative research with children. Furthermore, the methods used to validate the classification system engaged children both as active participants and as experts in their own health.

The involvement of children and young people was a priority throughout this study, from the identification of the preliminary classification system, through to the interpretation of the qualitative validation interviews. This level of involvement is rarely employed in the development of classification systems for paediatric preference-based measures, such as the generic EQ-5D-Y and HUI2, or condition-specific measures such as those for atopic dermatitis and asthma [34-37] .

Interestingly, children who had not experienced dental pain severe enough to cause them to cry were unable to understand the relevance of this impact. The range of responses surrounding this item from a sample who all have diagnosed dental caries confirms previous research highlighting the variation in impacts that children can experience, and how many suffer no symptoms at all [38]. Furthermore, the association between the number of carious teeth and the impacts experienced is often not as linear as one might expect [38]. Nonetheless, it is important for a utility measure to contain an item such as crying, since this is an impact that is only experienced by those with the greatest severity of the condition. This item, alongside the item *'kept awake',* will play an important role in the formation of the worst health state possible within the valuation survey [39,40].

Many participants within this study lived in areas that were amongst the most deprived in England, which reflects the association between caries prevalence and deprivation [41,42]. One potential limitation of this study is that it included disproportionately more female participants than males. This does not reflect the wider population, where there is a trend for boys to have a slightly higher prevalence of caries than girls [30].

The clinical caries experience (dmft/DMFT) of participants in this study was much higher than the national average of 0.9 [30]. The prevalence of caries in 5-year-old children in Yorkshire and the Humber is known to be

18

greater than the national average (28.5% compared to 24.7% respectively), though this discrepancy is more likely to be explained by the recruitment of participants from a tertiary referral centre. These participants are likely to have been referred to the dental hospital due to the extent of their disease, and resulting symptoms. Whilst this could be considered a limitation of the study due to the lack of representativeness of the sample, it could be argued that those experiencing the impacts described in CARIES-QC would be the most appropriate sample to validate the classification system. Furthermore, this approach ensured that those experiencing the most severe, and perhaps less frequently encountered impacts (e.g. crying) were involved.

In conclusion, following child-centred modification as detailed above, the preliminary classification system can now be considered valid, since it has been derived taking into account Rasch analyses, classical psychometric analyses, PPI and developer input, clinical input (HJR, HDR, ZM, FG) as well as involvement of children with dental caries. The five-item classification system is now suitable for use in a valuation survey with children and young people. This will facilitate determination of QALYs for children with caries, to better inform decision-makers and commissioners regarding the cost-utility of interventions to improve children's oral health. Furthermore, the innovative methodology used to develop and validate this classification system can be used in the development of other preference-based measures.

References

1. Benjamin, R. M. (2010). Oral Health: The Silent Epidemic. *Public Health Reports, 125*(2), 158-159, doi:10.1177/003335491012500202.
2. Majewski, R. F., Snyder, C. W., & Bernat, J. E. (1988). Dental emergencies presenting to a children's hospital. *ASDC J Dent Child, 55*(5), 339-342.
3. Schuch, H. S., Dos Santos Costa, F., Torriani, D. D., Demarco, F. F., & Goettems, M. L. (2015). Oral health-related quality of life of schoolchildren: impact of clinical and psychosocial variables. *International Journal of Paediatric Dentistry, 25*(5), 358-365, doi:10.1111/ipd.12118.
4. Pau, A., Baxevanos, K., & Croucher, R. (2007). Family structure is associated with oral pain in 12-year-old Greek schoolchildren. *Int. J. Paediatr. Dent., 17*(5), 345-351, doi:10.1111/j.1365-263X.2007.00842.x.
5. Krisdapong, S., Sheiham, A., & Tsakos, G. (2009). Oral health-related quality of life of 12- and 15-year-old Thai children: findings from a national survey. *Community Dentistry & Oral Epidemiology, 37*(6), 509-517.
6. Miller, J., Vaughan-Williams, E., Furlong, R., & Harrison, L. (1982). Dental caries and children's weights. *Journal of Epidemiology and Community Health, 36*(1), 49, doi:10.1136/jech.36.1.49.
7. Acs, G., Shulman, R., Ng, M. W., & Chussid, S. (1999). The effect of dental rehabilitation on the body weight of children with early childhood caries. *Pediatr Dent, 21*(2), 109-113.
8. HSCIC (2016). Hospital Episode Statistics England 2015-2016. Health and Social Care Information Centre.
9. Hettiarachchi, R. M., Kularatna, S., Downes, M. J., Byrnes, J., Kroon, J., Lalloo, R., et al. (2017). The cost-effectiveness of oral health interventions: A systematic review of cost- utility analyses. *Community dentistry and oral epidemiology*, doi:10.1111/cdoe.12336.
10. Tonmukayakul, U., Calache, H., Clark, R., Wasiak, J., & Faggion, C. M. (2015). Systematic Review and Quality Appraisal of Economic Evaluation Publications in Dentistry. *Journal of dental research, 94*(10), 1348-1354, doi:10.1177/0022034515589958.

11. Rogers, H. J., Rodd, H. D., Vermaire, J. H., Stevens, K., Knapp, R., El Yousfi, S., et al. (2019). A systematic review of the quality and scope of economic evaluations in child oral health research. *BMC Oral Health, 19*(1), 132, doi:10.1186/s12903-019-0825-2.

12. Lord, J., Longworth, L., Singh, J., Onyimadu, O., Fricke, J., Bayliss, S., et al. (2015). Oral Health Guidance – Economic analysis of oral health promotion approaches for dental teams Birmingham & Brunel Consortium: External Assessment Centre.

13. NICE (2013). Guide to the methods of technology appraisal. London: National Institute of Health and Care Excellence.

14. Foster Page, L. A., Beckett, D. M., Cameron, C. M., & Thomson, W. M. (2015). Can the Child Health Utility 9D measure be useful in oral health research? *International Journal of Paediatric Dentistry, 25*(5), 349-357, doi:10.1111/ipd.12177.

15. Gilchrist, F., Rodd, H. D., Deery, C., & Marshman, Z. (2018). Development and evaluation of CARIES-QC: a caries-specific measure of quality of life for children. *BMC Oral Health, 18*(1), 202, doi:10.1186/s12903-018-0662-8.

16. Gilchrist, F., Rodd, H., Deery, C., & Marshman, Z. (2014). Assessment of the quality of measures of child oral health-related quality of life. *Acta Veterinaria Scandinavica, 14*(1), 40, doi:10.1186/1472-6831-14-40.

17. McCabe, C., Stevens, K., Roberts, J., & Brazier, J. (2005). Health state values for the HUI 2 descriptive system: results from a UK survey. *Health Economics, 14*(3), 231-244, doi:10.1002/hec.925.

18. Young, T. A., Yang, Y., Brazier, J. E., & Tsuchiya, A. (2011). The Use of Rasch Analysis in Reducing a Large Condition-Specific Instrument for Preference Valuation: The Case of Moving from AQLQ to AQL-5D. *Medical Decision Making, 31*(1), 195-210, doi:10.1177/0272989X10364846.

19. Mulhern, B., Rowen, D., Brazier, J., Smith, S., Romeo, R., Tait, R., et al. (2013). Development of DEMQOL-U and DEMQOL-PROXY-U: generation of preference-based indices from DEMQOL and DEMQOL-PROXY for use in economic evaluation. *Health Technol Assess, 17*(5), v-xv, 1-140, doi:10.3310/hta17050.

20. King, M. T., Costa, D. S., Aaronson, N. K., Brazier, J. E., Cella, D. F., Fayers, P. M., et al. (2016). QLU-C10D: a health state classification system for a multi-attribute utility measure based on the EORTC QLQ-C30. *Qual Life Res, 25*(3), 625-636, doi:10.1007/s11136-015-1217-y.

21. Mulhern, B., Labeit, A., Rowen, D., Knowles, E., Meadows, K., Elliott, J., et al. (2017). Developing preference-based measures for diabetes: DHP-3D and DHP-5D. *Diabet Med, 34*(9), 1264-1275, doi:10.1111/dme.13377.

22. Marshman, Z., Gupta, E., Baker, S. R., Robinson, P. G., Owens, J., Rodd, H. D., et al. (2015). Seen and heard: towards child participation in dental research. *Int J Paediatr Dent, 25*(5), 375-382, doi:10.1111/ipd.12179.

23. Index of Multiple Deprivation (2015). Index of Multiple Deprivation. http://opendatacommunities.org/resource?uri=http%3A%2F%2Fopendatacommunities.org%2Fdata%2Fsocietal-wellbeing%2Fdeprivation%2Fimd-score-2010. Accessed September 2018.

24. Klein, H., Palmer, C., & Knutson, J. (1938). Studies on dental caries I: dental status and needs of elementary school children. *Public Health Reports, 53*, 751-765.

25. Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., et al. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology, 60*(1), 34-42.

26. Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.

27. Schafer, J. L. (1999). Multiple imputation: a primer. *Statistical Methods in Medical Research, 8*(1), 3-15, doi:10.1177/096228029900800102.

28. Gilchrist, F. (2015). Development of a child-centred, caries-specific measure of oral health-related quality of life. University of Sheffield, 2015.

29. Willis, G. B. (2005). *Cognitive interviewing [electronic resource] : a tool for improving questionnaire design*. Thousand Oaks, Calif. London : SAGE, c2005.

30. Pitts, N., Chadwick, B., & Anderson, T. (2015). Children's Dental Health Survey 2013. Report 2: Dental Disease and Damage in Children, England, Wales and Northern Ireland. Health and Social Care Information Centre.

31. Klein, H., & Palmer, C. E. (1940). COMMUNITY ECONOMIC STATUS AND THE DENTAL PROBLEM OF SCHOOL CHILDREN. [Article]. *Public Health Reports, 55*(5), 187-205, doi:10.2307/4583167.

32. Christensen, K., Makransky, G., & Horton, M. (2016). Critical Values for Yens Q3: Identification of Local Dependence in the Rasch Model Using Residual Correlations. *Applied Psychological Measurement, 41*, doi:10.1177/0146621616677520.

33. Gilchrist, F., Marshman, Z., Deery, C., & Rodd, H. D. (2015). The impact of dental caries on children and young people: what they have to say? *International Journal of Paediatric Dentistry, 25*(5), 327-338, doi:10.1111/ipd.12186.

34. Torrance, W. G., Feeny, H. D., Furlong, J. W., Barr, D. R., Zhang, D. Y., & Wang, D. Q. (1996). Multiattribute Utility Function for a Comprehensive Health Status Classification System: Health Utilities Index Mark 2. *Medical Care, 34*(7), 702-722, doi:10.1097/00005650-199607000-00004.

35. Wille, N., Badia, X., Bonsel, G., Burström, K., Cavrini, G., Devlin, N., et al. (2010). Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Quality of Life Research, 19*(6), 875-886, doi:10.1007/s11136-010-9648-y.

36. Stevens, K. J., Brazier, J. E., McKenna, S. P., Doward, L. C., & Cork, M. J. (2005). The development of a preference-based measure of health in children with atopic dermatitis. *British Journal of Dermatology, 153*(2), 372-377, doi:10.1111/j.1365-2133.2005.06736.x.

37. Chiou, C.-F., Weaver, M. R., Bell, M. A., Lee, T. A., & Krieger, J. W. (2005). Development of the multi-attribute Pediatric Asthma Health Outcome Measure (PAHOM). *International Journal for Quality in Health Care, 17*(1), 23-30, doi:10.1093/intqhc/mzh086.

38. Tickle, M., Milsom, K., King, D., Kearney-Mitchell, P., & Blinkhorn, A. (2002). The fate of the carious primary teeth of children who regularly attend the general dental service. *British Dental Journal, 192*(4), 219, doi:10.1038/sj.bdj.4801338.

39. Ratcliffe, J., Chen, G., Stevens, K., Bradley, S., Couzner, L., Brazier, J., et al. (2015). Valuing Child Health Utility 9D Health States with Young Adults: Insights from a Time Trade Off Study.

40. Ratcliffe, J., Couzner, L., Flynn, T., Sawyer, M., Stevens, K., Brazier, J., et al. (2011). Valuing child health utility 9D health states with a young adolescent sample. *Applied Health Economics and Health Policy, 9*(1), 15-27, doi:10.2165/11536960-000000000-00000.

41. Schwendicke, F., Dörfer, C. E., Schlattmann, P., Page, L. F., Thomson, W. M., & Paris, S. (2015). Socioeconomic Inequality and Caries: A Systematic Review and Meta-Analysis. (Vol. 94, pp. 10-18). Los Angeles, CA.

42. Slade, G. D., & Sanders, A. E. (2017). Two decades of persisting income-disparities in dental caries among U.S. children and adolescents. *Journal of Public Health Dentistry*, doi:10.1111/jphd.12261.