CSEF

Centre for Studies in Economics and Finance

WORKING PAPER NO. 508

Using Animation to self-report health: a Randomized Experiment with Children

Guerriero C., Abrines Jaume N., Diaz-Ordaz K., Brown K.,
Wray J. Ashworth J., Abbiss M., Cairns J.

September 2018

University of Naples Federico II

University of Salerno

Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS – UNIVERSITY OF NAPLES
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unina.it

CSEF
Centre for Studies in Economics and Finance

WORKING PAPER NO. 508

Using Animation to Self-report Health: a Randomized Experiment with Children

Guerriero C.[1], Abrines Jaume N.[2], Diaz-Ordaz K.[3], Brown K.[4], Wray J.[3], Ashworth J.[5], Abbiss M.[5], Cairns J.[2]

Abstract
The Child Health Utility-9D (CHU-9D) is the only generic preference based measure specifically developed to elicit health-related quality of life directly from children aged 7 to 11 years. The aim of this study was to investigate whether the use of animation on a touch screen device (tablet) is a better way of collecting health status information from children aged 4 to 14 years compared to a traditional paper questionnaire. The specific research questions were firstly, do young children (4 to 7 years) find an animated questionnaire easier to understand; secondly, independent of the child's age, is completion of an animated questionnaire easier for sick children in hospital settings; and thirdly, do children's preferences for the different versions of the questionnaire vary by the age of the child. Using a balanced cross-over trial we administered different versions of the CHU-9D to 221 healthy children in a school setting and 217 children with health problems in a hospital setting. The study tested five versions of the CHU-9D questionnaire: paper text, tablet text, tablet image, paper image and tablet animation. Our results indicate that the majority of the youngest children aged 4-7 years found the CHU-9D questions easy to answer independent of the type of questionnaire administered. Amongst children aged 7-14 with health problems the type of questionnaire was found to influence understanding. Children aged 7 to 11 years found tablet image and animation easier compared to text questionnaires while the oldest children in hospital found text based questionnaires easier compared to image and animation. Children in all three age groups preferred animation on a tablet to other methods of assessment. Our results highlight the potential for using an Animated Preference Based Measure to the health of children as young as 4 years.

Keywords: Children health valuation; Animation; self-reported health state measure; CHU-9D

JEL Classification: I10, I12,I31

Acknowledgments: The studies described in this work would have never been possible without the attention and enthusiasm of the children and without the precious help of their teachers. The authors are very grateful to the Directors of the schools participating in the project: The Queensbridge Primary School and the Reepham High School.

[1] Università di Napoli Federico II and CSEF. Corresponding author. E-mail: carla.guerriero@unina.it

[2] Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK.

[3] London School of Hygiene and Tropical Medicine, London, UK.

[4] Institute of Cardiovascular Science, University College London, London, UK.

[5] Animation, Royal College of Art, London, UK.

Table of contents

1. Introduction

Self-reported health state measures are an important input to economic evaluations of clinical interventions. Compared to other health benefit measures (e.g. life years saved), Quality Adjusted Life Years (QALYs) allow length of life and health status to be combined in a single outcome measure, as a result, national and international health economics guidelines currently recommend its use to assess the cost-effectiveness of different health care interventions [1].

Several preference-based measures (PBM) are available with which to estimate QALYs in adults. The EQ-5D is the most widely used health state description for Cost-Utility Analysis in the UK [2]. EQ-5D data are obtained by asking questions about five dimensions of a respondent's current health (mobility, self-care, usual activities, anxiety and depression, pain and discomfort). Despite its many advantages, the EQ-5D is inappropriate for children because it was created taking into account adults' views regarding the relevant dimensions of health, secondly, the wording and formats used to elicit preference based measures are tailored to adult populations [3] [4]. Thirdly, the values to be associated with each health state have been estimated from adult general population samples [4].

The assessment of health status in children has long been neglected because of the many inherent challenges of measuring health status in these age groups. Using adult PBMs, studies targeting children and adolescents have valued health status indirectly through teachers, parents and medical experts [5]. However, there is mounting evidence that adults are not good proxies for children and that children themselves are the best judges of their own physical and psychological well-being [6]. Results of a study conducted with a large sample of parents and children aged 8-11 years, for instance, suggest that children's health-related quality of life scores are less extreme than those expressed by parents on their behalf. The study also suggests that parents tend to underestimate the prevalence and the emotional impact of health status problems of their children [7].

According to a review conducted by Griebsch et al. the main reason for lack of child-reported health states is that the majority of generic measures are derived from adult populations[8]. Developed for children aged 7-11, the Child Health Utility-9D (CHU-9D) is the first pediatric preference-based instrument adopting a bottom-up approach [9]. Instead of using literature and experts' opinion, as in previous PBMs, the CHU-9D

has been developed using one-to-one interviews conducted with children in order to identify health dimensions relevant to children and to assess children's ability to understand the questions [4]. Another advantage of the CHU-9D is that all of the dimensions of the questionnaire are severity based (while the HUI2, for example, contains a mix of severity and frequency). Designed for self-completion the CHU-9D generic PBM assesses nine health dimensions each comprising five levels that are relevant to children and young adolescents: worried, sad, pain, tired, annoyed, schoolwork/homework, sleep, daily routine and activities. According to Stevens [10] children using the CHU-9D can easily understand and describe their health, however, they have difficulty remembering when a particular problem occurred.

Despite the above-mentioned advantages that make CHU-9D the best PBM currently available for children, there are two main limitations: as it is designed for self-completion, children are required to be able to read and understand the questions and therefore it cannot be administered to many children (e.g. younger than 7 years) without the assistance of an adult. In addition, especially for severely ill children, reading and completing a written questionnaire can be tiring and sometimes physically impossible. Even if children are old enough to understand the written CHU-9D questions, when they are ill they may have reduced concentration and/or they may be too tired to complete the questionnaire by themselves [4]. For instance, in the pilot study conducted by Stevens to assess the psychometric performance of the CHU-9D, 95% of children in the school sample read the instructions themselves while 49% of the clinical sample asked for help to understand the instructions because they were too tired or felt too unwell [10].

This research project aimed to investigate whether the use of animation on a touch screen device (tablet) is a better way of collecting health status information from children compared to a traditional paper questionnaire. Compared to written descriptions, the use of graphic pathographies - illness narratives in graphic form - to display health conditions has been shown to be a useful way of providing detailed insights about the various aspects of a disease and reducing patient anxiety associated with the disease[11]. Improvements in computer technologies and falling production costs make animation an attractive and relatively inexpensive interactive tool to foster children's cognitive abilities. Animation can be used to present temporal change through movement compared to static or still images and is practical when information is inherently dynamic such as with biological processes. Findings from studies of cognitive

theories of multimedia learning largely support the idea that graphic presentations increase learning and understanding. For instance, a recent study conducted with children with autism spectrum disorder shows that children of all ages strongly prefer animation to other media forms (e.g. sport programs, music shows etc.) [12].

During the first part of the project an interdisciplinary group of researchers (child psychologists, a paediatrician, animation filmmakers, and health economists) developed with children of different ages, an interactive application presented on a touch screen device using an animated character named "Mix" (Figure 1). The application was designed to be suitable for young respondents of different ages, gender and ethnicity. Further details on how the APBM was developed can be found in Abrines Jaume et al. 2015[13].

2. Methods

2.1. Study aim and research questions

The aim of this research was to establish whether animation could be used to facilitate the self-reporting by children of their health status. Three complementary research questions were addressed:

(1) Does animation improve understanding of the CHU-9D questions among children aged 4 to 14 years, and particularly among the youngest children (4-7 years)?

(2) Does the use of animation have a differential impact on sick children's ability to understand and complete the CHU-9D questionnaire compared to that of healthy school children?

(3) Do children prefer to report their health state using animation, and do their preferences differ by age and whether they are sick or healthy?

Compared to a standard text CHU-9D, children receiving the APBM and still images on a tablet do not need to read the questions, but just look at five characters depicting the different dimension levels and tap on the screen to indicate the one that is most similar to how they feel (See Figure 2 for an example of still Image). Hence, we expected that both methods, APBM and Tablet Images, would be easier for children to use and hence have greater acceptability, especially in the context of health problems and therefore our third aim was to investigate which method children preferred to report their health status.

2.2. Health state assessment methods

When testing the specific hypotheses of the study there are a number of differences between the APBM and the standard text-based questionnaire that must be considered: the effects of the use of animation itself; the description of dimensions using still images versus words and the use of a touch screen device for data entry. In order to disentangle the effect of using animation from the other potential differences five versions of the CHU-9D questionnaire were tested in this study (Figure 3):

1. Animated presentation with touch screen data entry (A)
2. Still graphic images with touch screen data entry (B)
3. Hard copy questionnaire with graphic images (C)
4. Touch screen responses to text-based questionnaire (D)
5. Hard copy text-based questionnaire (E)

Given the five modes of questionnaire administration there are ten potential contrasts, one of which compares the animated questionnaire to text-based hard copy data collection, and six that enable the impact of a single difference between questionnaires to be tested. The three remaining contrasts are potentially less interesting since they do not involve the A:E contrast of primary interest nor do they permit the effect of a single difference to be identified.

As indicated in Figure 4 the sample comprised three age groups: children aged 4-7 years; children aged 7-11 and older children aged 11-14 years. The age groups for school and hospital children corresponded to classes which follow school curriculum years and hence there was some overlap of ages at the extremes of age within each group. In order not to overburden participants it was decided to present only three of the five forms of the questionnaire to any one participant. In each case the order of questionnaires was randomised.

Group one, aged 4-7 years

It was anticipated that many of the youngest age group, children aged 4-7, would be unable to engage in text-based tasks and therefore the use of text was avoided as much as possible. Only one task (E) involved a text-based questionnaire. This was included since the comparison of A with E was of primary interest for the project. The effects of animation were assessed by comparing A with B.

Group two, aged 7-11 years

Children in the second age group (aged 7-11 years) are better readers and all the relevant comparisons were tested with this age group (see Figure 1). These children were randomly allocated to receive either A, B and E or to receive B, C and D. The responses from the latter group permit investigation of the effect of using a still image and the effect of using a tablet, while the former group provide data relevant to the primary comparison of interest between tablet animation and paper text.

Group three, aged 11-14 years

Children in the oldest age group (11-14 years) were asked to complete versions, A, D and E. It was expected that the majority of children in this age group would have no difficulties in completing more than one text based measure. The comparison of A with E explores whether animation is the preferred format even among more accomplished readers. The comparison of D with E isolates the impact of Tablet presentation, while the comparison of A with D investigates whether or not children prefer text to animation when both versions of the questionnaire are presented using a touch screen device.

2.3. Study Setting, inclusion criteria and procedures

The study population consisted of both sick and healthy children aged between 4 and 14 years from both school and hospital settings. Healthy children were recruited from two primary schools admitting children aged between 4 and 11 years in London and one secondary school admitting children aged from 11 years and including those up to 14 years. All the schools included reflect the diverse population of inner London.

Children with health problems were recruited from one tertiary paediatric hospital in London. Both inpatient wards and day case areas were included and children had a range of health conditions.

In terms of inclusion and exclusion criteria, as well as meeting the age specification, children were required to speak fluent English. Children with learning disabilities were excluded in order to increase the homogeneity of the sample.

For the healthy children recruited from a school setting, invitation letters and information sheets were sent from the school to parents of children who met the inclusion criteria. Researchers went to the school and

explained what the research project was about to children whose parents provided signed consent. Children who wanted to take part gave written assent using age appropriate forms created specifically for this study. Different information sheets and assent forms were designed for each age group. For the children with health problems recruited within a hospital setting, each individual child and their parent who were approached for participation were given age-appropriate information. Parents were asked to provide written informed consent for their child's participation in the study and children provided written assent for their own participation.

Once assent and consent forms were signed, the research task was administered by the researcher by opening the sealed envelope randomly allocated to the child, containing their ID number, the tasks they had to do and the order in which they had to be administered.

2.4. Design, Outcomes and Analysis

The study was designed as a balanced cross-over trial, uniform within sequences and task, across age groups. The tasks were consecutive and a distraction activity was always presented between the experimental conditions. We assumed that there was no significant carry-over effect. Each child was presented with three experimental conditions in random order but the number of experimental conditions tested in the study varied between age groups (three experimental conditions in the youngest and oldest age groups; five in the middle age group). Randomisation of treatment sequences within each age group was stratified by whether the child was healthy (school setting) or unwell (hospital setting).

A previous study conducted by Canaway and Frew (2013) was used to determine the sample size for the two younger age groups (4-7 and 7-11 year-old)[14]. In Canaway's and Frew's (2013) study interviewer ratings were obtained of the understanding of children (aged 6-7 years) completing the CHU-9D paper based questionnaire. The power analysis calculations were based on these study results and accounted for the younger ages of our respondents. Assuming 80% power, a significance level of 0.05 and an attrition rate of 15%, we estimated the minimum number of subjects required for each questionnaire type to test the primary outcome, children's understanding, by age group and setting.

The school children's understanding was assessed and their preferences collected by one of the three researchers administering the questionnaires. Only one researcher administered the questionnaires to children with health problems because randomization was slower in the hospital setting. Researchers rated the child's understanding using a 9 category classification subsequently transformed for the analysis into a dichotomous outcome measure indicating either good or poor understanding. Similarly, children were asked to rate each version of the questionnaire using a 5 point Likert scale (very poor, poor, fair, good and excellent). This measure was also transformed into a binary variable indicating whether the child found the questionnaire either easy or difficult to understand (good and excellent rating were grouped together indicating that the child found the questionnaire easy to understand).

At the end of the questionnaire children were asked to indicate which version of the CHU-9D questionnaire they preferred.

3. Primary outcomes

In the youngest age group (4-7 years), the child's understanding was evaluated using both the researcher's and the child's rating as co-primary outcomes. For the middle and older age groups the child's rating was the sole primary outcome used to indicate how difficult the respondent found the different questionnaires. The secondary outcome was which questionnaire was preferred by each age group. As with understanding, a binary outcome measure (preferred; not preferred) was used.

Statistical Analysis

The outcomes were analysed using a random subject effect logistic mixed-effect model. The model used was:

$$logit\left(\pi_{ij}\right) = \beta_0 + t_{kj} + hospital_j + Age + Gender + p_{ij} + u_j$$

Where: β_0 is the fixed intercept; $hospital_j$ is a binary variable indicating whether the child was recruited in a school vs. hospital setting; t_{kj} is the task (k=1,2,3,4,5) by subject j; *Age* is the age variable, and has been centered at 6, 9 and 13 for each of the three age groups respectively; *Gender* is a binary variable controlling for gender; p_{ij} is the fixed effect associated with the order of the task i (i = 1, 2, 3); u_j is the subject effect

(random effect associated with subject *j*). We assume $u_j \sim N(0, {\sigma_u}^2)$. Secondary analyses were also conducted to test whether there were differences in understanding and preferences by subgroup by adding an interaction term between hospital and treatment.

4. Results

4.1. Descriptive statistics

A total of 438 children were approached and consented to participate in the study. However, it was only possible to obtain complete information from 391 children because 47 children did not complete all the three tasks. A Table reporting a description of missing data and why they are missing is reported in the Appendix.

Table 1. Characteristics of the sample groups within the study population

Group one, 4-7 years		**Group two, 7-11 years**		**Group three, 11-14 years**	
4 years	30 (18%)	**7 years**	17 (9%)	**11 years**	25 (9%)
5 years	63 (39%)	**8 years**	58 (31%)	**12 years**	34 (31%)
6 years	43 (26%)	**9 years**	38 (20%)	**13 years**	22 (20%)
7 years	28 (17%)	**10 years**	46 (25%)	**14 years**	5 (25%)
		11 years	29 (15%)		
Female : Male	83 : 81		82 : 106		44 : 42
School : Hospital	83 : 81		96 : 92		42 : 44
Total	164		188		86
Who read the questionnaire?					
Child	6 (4%)		131 (70%)		75 (87%)
Researcher	138 (84%)		18 (10%		3 (4%)
Jointly read by child & researcher	11 (7%)		23 (11%)		2 (2%)
Missing	9 (5%)		16 (9%)		6 (7%)

Children with health problems were more likely to ask the help of the researchers to read the questionnaire. Researchers were the sole reader for 70, 12 and 2 hospital children in the three different age groups.

Tables 1 describes the three sample groups of children included in the study. Overall, 173 children in the youngest age group (aged 4-7 years) were approached and all assented to participate in the study. Table 1 provides information about age, gender, health group and who read the questionnaire. As expected in this age group the majority of children asked the interviewer to read the paper questionnaires (84%). The second group (aged 7-11 years) was composed of 188 children. Compared to the younger age group, only a small proportion of children asked the researcher to read the text questionnaires, this being 11% who asked for help in reading the text material. The oldest age group was composed of 86 children aged between 11 and 14 years, and as expected, only 4% of children (all of them from hospital settings) asked the researcher to read the questionnaire. Across all age groups, the sample was age and gender balanced.

4.2. Questions

1) Does APBM enhance understanding of CHU-9D dimensions particularly amongst the youngest children?

The majority of children (69.59%) in the youngest age group found the CHU-9D questions either easy or very easy to answer. As expected, this proportion rose to 71% and 82% among the intermediate and oldest children.

Table 2 shows that for children aged 4-7 years there were no differences in understanding between the different formats of the questionnaire. There was, however, weak evidence that for every year of age the odds of the child reporting the tasks to be easy increased. Again, in the youngest age group when the outcome considered was the researcher's rating of the children's understanding, the relationship between increasing age and the odds of finding the task easy was preserved and was statistically stronger (Table 3). Furthermore, rating of the task as easy was more likely when presented via Tablet Still Image or APBM versus Paper Text. None of the other factors considered (task order, gender, hospital or school setting) had a statistical relationship to either outcome.

Analysis of the children's rating of understanding in the intermediate and oldest age groups (groups two and three, as shown in Table 2) also indicated that there were no differences detected in this rating between the types of questionnaires. In the intermediate age group there was also a significant increase in the odds ratio of finding the questionnaire easy with each year of age.

2) Does APBM enhance the understanding of CHU-9D dimensions amongst children with health problems?

Tables 2-3 show the results of an investigation of whether there is a treatment-by-condition interaction, that is, whether the type of questionnaire had a different effect depending on whether the child was a healthy participant assessed in school or a participant with health problems assessed in hospital. As presented, independent of the outcome variable used for rating (researcher's vs. child's) none of the coefficients for the interactions are significant within the youngest age group. Absence of a significant difference may be due to the children not having to read the questionnaire themselves.

In the intermediate age group, however, the children with health problems completing the APBM and Tablet still image questionnaires were more likely to rate the task as easy compared to text-based questionnaires (both paper and tablet text; Table 2). In the oldest age group, children found it easier to complete text questionnaires (both Tablet and Paper text) compared to APBM. In both intermediate and older age groups, the children with health problems were more likely than the healthy children from a school setting to rate the tasks as easy, possibly because children with health problems are more experienced in answering these types of questions.

Table 2: Random subject-effect logistic model. Outcome variable: child's rating of understanding (binary Easy=1/Difficult=0) by setting and type of task, by age group.

Variable	**OR**	**SE**	**CI**
Youngest Age Group			
Paper Text vs. Tablet Still Image	0.91	0.37	0.41-2.01
Tablet Animation vs. Tablet Still Image	0.47	0.19	0.22-1.03
Hospital vs. School	0.88	0.43	0.34-2.30
Hospital & Paper Text vs. Hospital & Tablet Still Image	0.53	0.30	0.17-1.63
Hospital & Tablet Animation vs. Hospital & Tablet Still Image	2.09	1.20	0.68-6.41
Age	1.38	0.25	0.97-1.98
Gender	1.45	0.51	0.72-2.91
Order of Task 1 vs. 2	1.18	0.33	0.68-2.04
Order of Task 1 vs. 3	1.41	0.40	0.81-2.46
Constant	2.15	1.33	0.64-7.28
Intermediate Age Group			
Paper Image vs. Tablet Still Image	2.76	1.50	0.95-8.00

Paper Text vs. Tablet Still Image	1.42	0.72	0.52-3.83
Tablet Animation vs. Tablet Still Image	0.81	0.40	0.30-2.15
Tablet Text vs. Tablet Still Image	2.37	1.26	0.84-6.73
Hospital vs. School	7.58***	4.46	2.39-24.02
Hospital and Paper Still Image vs. Tablet Still Image	0.26	0.22	0.05-1.34
Hospital & Paper Text vs. Hospital & Tablet Still Image	0.20*	0.15	0.04-0.90
Hospital & Tablet Animation vs. Hospital & Tablet Still Image	0.50	0.38	0.11-2.24
Hospital & Tablet Text vs. Hospital & Tablet Still Image	0.13*	0.10	0.02-0.63
Age	1.48*	0.26	1.05-2.10
Gender	0.96	0.41	0.69-1.45
Order of Task 1 vs. 2	1.48	0.45	0.81-2.68
Order of Task 1 vs. 3	1.04	0.31	0.58-1.85
Constant	1.58	1.25	0.34-7.42
Oldest Age Group			
Tablet Animation vs. Paper Text	1.54	0.96	0.45-5.22
Tablet Text vs. Paper Text	1.54	0.95	0.46-5.18
Hospital vs. School	5.19*	3.95	1.17-23.09
Hospital & Tablet Animation vs. Hospital & Paper Text	0.12*	0.11	0.02-0.75
Hospital & Tablet Text vs. Hospital & Paper Text	0.24	0.23	0.04-1.59
Age	1.18	0.31	0.71-1.98
Gender	0.52	0.25	0.20-1.35
Order of Task 1 vs. 2	1.41	0.63	0.60-3.38
Order of Task 1 vs. 3	1.93	0.89	0.78-4.75
Constant	10.61*	9.83	1.72-65.23

Notes: OR: Odds Ratio; SE: Standard Error; CI: 95% Confidence Interval *** significant at $p<0.001$;** significant at $p<0.005$; * significant at $p<0.05$

Table 3. Random subject-effect logistic model. Outcome variable: researcher's rating of understanding (binary Good understanding =1/Poor understanding =0) for children in the youngest age group.

Variable	OR	SE	CI
Paper Text vs. Tablet Still Image	0.70	0.44	0.20-2.40
Tablet Animation vs. Tablet Still Image	1.08	0.71	0.30-3.89
Hospital vs. School	0.86	0.72	0.17-4.43
Hospital & Paper Text vs. Hospital & Tablet Still Image	0.20	0.17	0.04-1.08
Hospital & Tablet Animation vs. Hospital & Tablet Still Image	2.30	2.12	0.38-13.99
Age	4.99***	2.04	2.24-11.13
Gender	0.66	0.42	0.20-2.33
Order of Task 1 vs. 2	1.57	0.69	0.66-3.71
Order of Task 1 vs. 3	1.49	0.65	0.63-3.51
Constant	0.01*	0.02	0.001-0.69

Notes: OR: Odds Ratio; SE: Standard Error; CI: 95% Confidence Interval *** significant at $p<0.001$;** significant at $p<0.005$; * significant at $p<0.05$

3) Do children prefer APBM as a means to assess their health state?

Table 4 shows the results of the preference analysis. Across all the age groups APBM was the most preferred type of questionnaire, however, there are interesting differences by age group between the other possible formats of the CHU-9D. Children in the youngest age group preferred APBM versus both Tablet Still Image and Paper Text. The preference appears to be for animation rather than simply wanting to use a tablet since no difference was found between Paper Text and Tablet Still Image. In the intermediate age group, children preferred APBM to Tablet Still Image, but also preferred Tablet Text to Still Image though to a lesser extent than their preference for APBM. Among the oldest age group, children preferred method was APBM followed by Tablet Text with their least preferred method being Paper Text. The two older age groups appear to be expressing general preference for questions supported by a mobile device rather than paper.

Table 4: Random subject-effect logistic models of preferences by age group.

Variable	OR	SE	CI
Youngest Age Group			
Animation vs. Paper Text	17.36***	5.32	9.52-31.67
Tablet Still Image vs. Paper Text	1.87	0.61	0.98-3.56
Hospital vs. School	0.94	0.22	0.59-1.50
Age	1.08	0.13	0.86-1.36
Gender	0.98	0.22	0.59-1.49
Order of Task 1 vs. 2	1.40	0.40	0.80-2.43
Order of Task 1 vs. 3	1.25	0.35	0.71-2.18
Constant	0.11***	0.05	0.04-0.26
Intermediate Age Group			
Paper Image vs. Tablet Image	0.57	0.19	0.29-1.08
Paper Text vs. Tablet Image	0.67	0.21	0.36-1.24
Animation vs. Tablet Image	7.92***	2.26	4.52-13.90
Tablet Text vs. Tablet Image	1.86*	0.50	1.09-3.16
Hospital vs. School	0.98	0.19	0.66-1.44
Age	0.98	0.08	0.84-1.15
Gender	0.93	0.19	0.63-1.38
Order of Task 1 vs. 2	0.91	0.22	0.56-1.46
Order of Task 1 vs. 3	1.05	0.81	0.66-1.69
Constant	0.38*	0.10	0.10-.52
Oldest Age Group			
Animation vs. Paper Text	23.78***	13.21	7.99-70.76
Tablet Text vs. Paper Text	13.54***	7.56	4.53-40.47
Hospital vs. School	1.00	0.32	0.53-1.85
Age	0.94	0.16	0.68-1.33
Gender	1.09	0.34	0.59-2.00
Order of Task 1. vs. 2	1.15	0.41	0.57-2.32
Order of Task 1 vs. 3	1.19	0.44	0.58-2.45
Constant	0.04**	0.02	0.01-0.15

Notes: OR: Odds Ratio; SE: Standard Error; CI: 95% Confidence Interval *** significant at $p<0.001$;** significant at $p<0.005$; * significant at $p<0.05$.

5. Discussion

To the best of our knowledge, this is the first study investigating whether the use of animation on a tablet device is a better means of asking children aged 4-14 years about their health state. The results of the study and their implications may be summarised as follows: in the youngest age group consisting of 4-7 year olds, the majority of children asked the researcher to read all textual information across the measures used. Compared to the healthy children participating at school, the children with health problems were also more likely to require help from the researcher with reading textual material. A strong advantage of APBM in this respect is that the use of graphic image depictions of the various health domains enables the use of only one word to describe the health state and as a result, the amount of reading and the requested assistance is significantly reduced compared to text-based questionnaires. This feature enabled the younger and hospitalised children to participate in the evaluation more independently of adult supervision, which may be preferable for them and reduces the influence of third parties upon their responses. This may be particularly important for hospitalised children who frequently lose some of their ability to be independent due to their condition or requirement for invasive treatment. A novel finding of this study is that even children as young as four years can understand the CHU-9D questions and report their health status directly using the nine dimensions of the questionnaire, although as expected the ability to understand the questionnaire improves with age. When the ratings of the children's understanding was considered within the youngest age group of 4 to 7 year olds, the researcher found the APBM and Tablet to be easier for the children than paper questionnaires, however, the children themselves did not indicate any difference between measures. Therefore, a further advantage of the APBM is that visually presented questions may be easier for younger children (under the age of 7 years) to understand. This is advantageous because this age group has been usually excluded from previous studies [4,9,10].

The ability of children with health problems to understand the questionnaire was influenced by the format of the questionnaire and the easiest format varied by age group. Within the youngest age group (4-7 year olds), there was no difference in understanding of different formats. In the intermediate age group (7-11 year olds), amongst those with health problems the APBM and Tablet Still Image were found to be easier to complete than text based questionnaires. Hence, within this age group the use of a hand held electronic device has the potential to improve access to health state evaluation in a hospital setting. Amongst the oldest age group (11-

14 year olds) the children with health problems found the text-based questionnaire easiest to understand. One possible explanation for this is that teenagers with chronic health conditions are likely to have previous experience in completing text-based questionnaires about their health [10].

Preference analyses suggested that the APBM was the preferred method across all of the three age groups. Following APBM, the youngest age group (4-7 years old) preferred Tablet Still Image to Text whilst the intermediate (7-11 years old) and older (11-14 year old) children preferred Tablet Text to both Tablet Still Image and, as expected, Paper Text. The preferences expressed for APBM are important because they may underpin greater engagement of children across the age range in the evaluation of their health state with future use of APBM. The engagement of children in the direct evaluation of their health is ultimately more likely to generate inclusive and high quality data that are pertinent to health care evaluations.

The focus of this paper is to analyse the impact of different formats on children's understanding and preferences. A future complementary study will report tests of validity for these formats (e.g. can APBM discriminate between groups with different health problems?) using additional data collected during the study. Similarly, the analysis reported here does not include those children who did not complete the questionnaire. A subsequent study is being conducted to analyse informative missingness using the data reported in the Appendix. A potential limitation of this study that may be addressed by future research is that it did not attempt to adjust for socio-economic status or for other factors reflecting the children's background such as ethnicity. Furthermore, non-English speaking children were not included in this stage of the research, although APBM might increase the likelihood of children who do not speak English being able to self-report because of the limited use of words. Children with health problems involved in this research had a range of conditions including renal, oncological, haematological and respiratory diseases, however the study did not attempt to adjust for the type or the duration of the condition, which may have influenced the results. Another limitation to consider is that although the study found that the three age groups preferred ABPM, this preference may disappear over repeated exposure, particularly in the case of older children.

Few studies have been conducted to elicit health state valuations directly from children. Nevertheless the evidence suggests that adults are not able to report children's physical and psychological well-being [6]. This study, for the first time, attempted to elicit health state directly from children as young as four years. It also tests whether animated preference based measures are easier to understand and preferred compared to other

formats of CHU-9D. In summary, this study indicates that it is possible to evaluate directly the health state of children as young as four years old and that the CHILDSPLA APBM has features that may enhance the direct participation of young and sick children in the evaluation of their health.

The results of this project would have never been possible without close collaboration between experts from (usually) non-communicating disciplines: Medicine, Economics, Psychology and Animation Art. The results of this interdisciplinary collaboration revealed the considerable potential for further interdisciplinary research, for example, further research on the elicitation of health state information from even younger children, and from adults and children with learning or sensorial disabilities. Given the difficulties for parents to understand how their children feel APBM may have potential as a family communication tool.

References

1. National Institute for Clinical Excellence (NICE) Updated guide to the methods of technology appraisal. 2008. 29-49.
2. Dolan, P. 1997. Modelling valuations for EuroQol Health States. Medical Care. 35, 1095-1108.
3. Eiser, C., Morse, R., 2001. Quality-of-life measures in chronic diseases of childhood. [5;4]. Health Technology Assessment. 5,4.
4. Stevens, K., 2009. Developing a descriptive system for a new preference-based measure of health-related quality of life for children. Qual Life Res. 18, 1105-13.
5. Ravens-Sieberer U., et al., 2006. Generic health-related quality-of-life assessment in children and adolescents: methodological considerations. Pharmacoeconomics. 24, 1199–1220.
6. Apajasalo, M., et al., 1996. Quality of life in pre-adolescence: a 17-dimensional health-related measure (17D). Qual Life Res. 5, 532-8.
7. Teunnissen NC., 1998. The proxy problem: child report versus parent report in health-related quality of life research. Qual Life Res. 7, 387-97,.
8. Griebsch, I., Coast, J., Brown, J., 2005. Quality-Adjusted Life-Years Lack Quality in Pediatric Care: A Critical Review of Published Cost-Utility Studies in Child Health. Paediatrics. . 115, e600-e614.
9. Stevens, K.J., 2008. Working with children to develop dimensions for a preference based generic paediatric health related quality of life measure. Health Economics and Decision Science Discussion Paper 08/04.
10. Stevens, K., 2011. *Assessing the performance of a new generic measure of health-related quality of life for children and refining it for use in health state valuation.* Appl Health Econ Health Pol. 9(3, 157-69.
11. Green, M.J., Myers, K.R., 2010. Graphic medicine: use of comics in medical education and patient care. BMJ. 340, c863.

12. Shane, H.C., Albert, P.D.J., 2008. Electronic Screen Media for Persons with Autism Spectrum Disorders: Results of a Survey. J Autism Dev Disord. 2008. 38, 1499-508.
13. Abrines Jaume, N., Abbiss, M., Wray, J., Ashworth, J,, Brown, K., Cairns, J., 2015. CHILDSPLA: a collaboration between children and researchers to design and animate health states. Child: Care, Health and Development. 41, 1140-51.
14. Canaway, A.G., Frew, E.J., 2013. Measuring preference-based quality of life in children aged 6-7 years: a comparison of the performance of the CHU-9D and EQ-5D-Y--the WAVES pilot study. Qual Life Res. 22, 173-83.

Figure 1: “Mix” the protagonist of our study

Figure 2: Example of Still Graphic Image on a Tablet device.

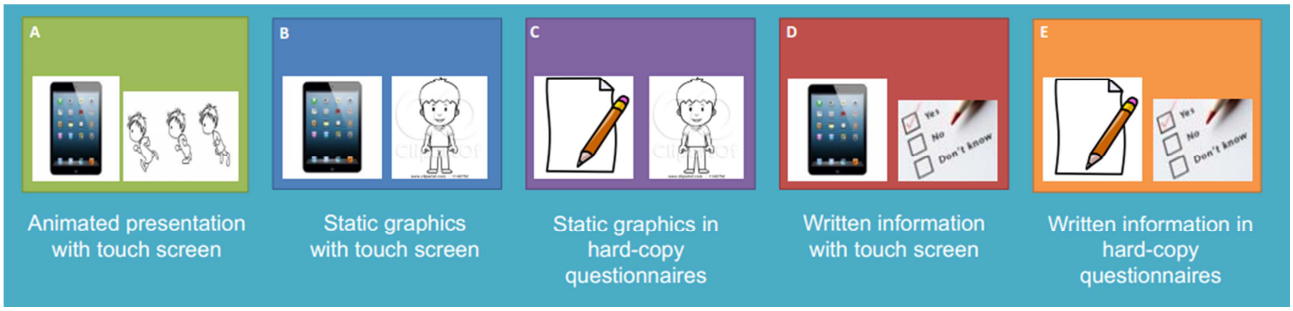

Figure 3. The formats of CHU-9D tested in our study

		A tablet animation	B tablet still image	C paper still image	D tablet text
E	paper text	Primary contrast 4-7/7-11/11-14	Combined effect 4-7/7-11	Effect of still image Not available	Effect of tablet 11-14
D	tablet text	Effect of animation 11-14	Effect of still image 7-11	Combined effect 7-11	
C	paper still image	Combined effect Not available	Effect of tablet 7-11		
B	tablet still image	Effect of animation 4-7/7-11			

Figure 4: Contrasts available between modes of questionnaire administration by age group