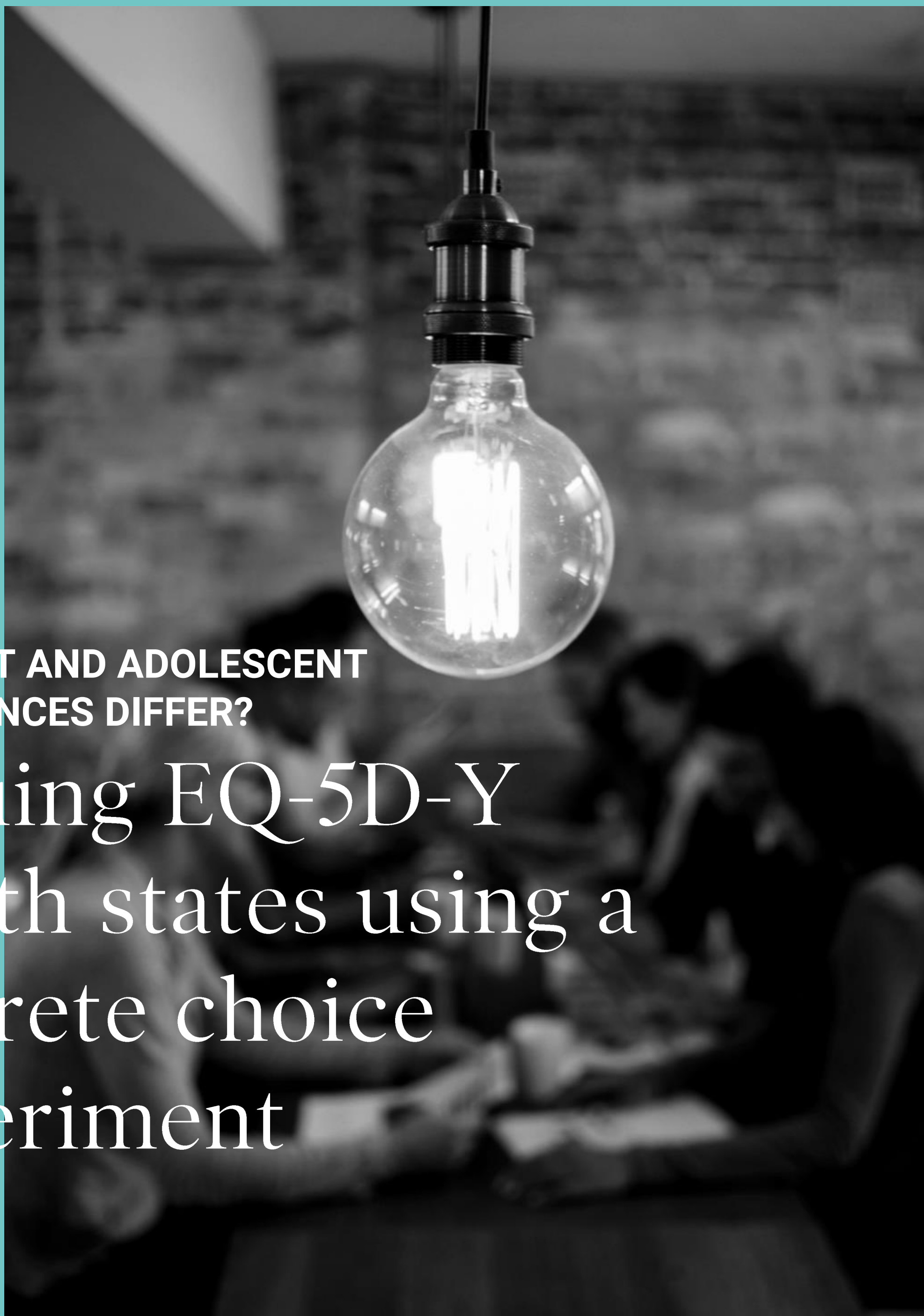
OHE



RESEARCH | REPORT
JULY 2019

DO ADULT AND ADOLESCENT PREFERENCES DIFFER?

Valuing EQ-5D-Y health states using a discrete choice experiment

David J. Mott
Koonal K. Shah
Juan Manuel Ramos-Goñi
Nancy J. Devlin
Oliver Rivero-Arias

JULY 2019

DO ADULT AND ADOLESCENT PREFERENCES DIFFER?

Valuing EQ-5D-Y health states using a discrete choice experiment

David J. Mott
Evidera, London, UK
Office of Health Economics, London, UK

Koonal K. Shah
Office of Health Economics, London, UK

Juan Manuel Ramos-Goñi
Axentiva Solutions, Tenerife, Spain
Executive Office, EuroQol Research Foundation, Rotterdam, The Netherlands

Nancy J. Devlin
Melbourne School of Population and Global Health, University of Melbourne, Australia
Office of Health Economics, London, UK

Oliver Rivero-Arias
National Perinatal Epidemiology Unit (NPEU), Nuffield Department of Population Health, University of Oxford, UK

OFFICE OF HEALTH ECONOMICS
RESEARCH

Please cite this report as:
Mott, D. J., Shah, Ramos-Goñi, J. M., Devlin, N.J., Rivero-Arias, O., 2019. *Valuing EQ-5D-Y health states using a discrete choice experiment: do adult and adolescent preferences differ?* OHE Research Paper. London: Office of Health Economics. Available at: https://www.ohe.org/publications/valuing-eq-5d-y-health-states-using-discrete-choice-experiment-do-adult-and-adolescent

Corresponding Author:
Oliver Rivero-Arias
oliver.rivero@npeu.ox.ac.uk

For further information please contact:

Professor Mireia Jofre-Bonet
Vice President and Head of Research, OHE
Professor in Economics at City, University of London

Tel +44 (0)207 747 8864
Email MJofre-Bonet@ohe.org

About Research Papers

OHE Research Papers are intended to provide information on and encourage discussion about a topic in advance of formal publication. Due to the requirements of the target journal for this manuscript, it has not undergone review by a member of OHE's Editorial Panel or any other form of formal peer review. Any views expressed are those of the authors and do not necessarily reflect the views or approval of the OHE's Editorial Panel or Research and Policy Committee, or its sponsors.

Once a version of the Research Paper's content is published in a peer review journal, that supersedes the Research Paper and readers are invited to cite the published version in preference to the original version.

Funding and Acknowledgements

Funding for this research was provided by a EuroQol Research Foundation grant. The views expressed in this paper are those of the authors, and are not necessarily those of the EuroQol Group. The authors are grateful to Rosalie Viney for helpful comments on an earlier draft

Table of Contents

OFFICE OF HEALTH ECONOMICS
RESEARCH

OHE

Abstract

Introduction: One of the challenges with generating an EQ-5D-Y value set is that traditional methods are cognitively demanding and may not be appropriate for younger individuals. However, asking adults to complete a valuation task from the perspective of a child/adolescent presents its own challenges. In this study we collected adolescent and adult responses to a discrete choice experiment (DCE) containing EQ-5D-Y health states in order to determine whether the two groups exhibit different preferences.

Methods: An online survey was designed containing a DCE, which comprised 15 pairwise choices. A sample of UK adults was asked to consider the health of a 10-year-old child when completing the tasks. In contrast, a sample of UK adolescents (11-17 years) received the same survey and completed the tasks considering their own health. Mixed logit models were estimated for both samples and comparisons were made.

Results: In total, 1,000 adults and 1,005 adolescents completed the survey. The relative importance of the levels attached greatest disutility to level 3 in pain/discomfort (PD3) followed by anxiety/depression (AD3) in both groups. The rank-order of other levels differed, including the third-worst level: mobility (MO3) for adolescents; and usual activities (UA3) for adults. Modelling results indicate that there are significant differences in preferences between the two samples.

Limitations: The perspective of the task differed between the two samples and therefore we cannot determine whether the differences are due to the perspective of the task or individuals' preferences.

Conclusions: Adolescents' preferences differ from those of adults taking the perspective of a child. As the adolescents were capable of completing the DCE, it is important to consider whether their preferences should be considered in decision-making.

1 Introduction

The EQ-5D-Y is a patient-reported outcome measure that was designed to measure the health-related quality of life of children and adolescents (Wille et al., 2010; Ravens-Sieberer et al., 2010). However, unlike the adult versions of the same instrument (EQ-5D-3L and EQ-5D-5L), currently no value sets exist such that responses to the EQ-5D-Y can be translated to health state utilities. This means that in economic evaluations of treatments aimed at younger populations it is not possible to estimate quality-adjusted life years (QALYs) based on EQ-5D-Y data, unless a value set for another instrument is used. As the EQ-5D-Y instrument contains three severity levels, it is typical for EQ-5D-3L value sets to be used in lieu of a value set for the EQ-5D-Y. However, evidence suggests that this is inappropriate because the two instruments are worded differently and the perspective used in an EQ-5D-Y valuation task may result in a value set with significantly different characteristics to an EQ-5D-3L value set (Kind et al., 2015; Kreimeier et al., 2018). It is therefore important that preference elicitation studies are conducted in order to generate a value set for the EQ-5D-Y instrument.

Whilst the development of value sets for the EQ-5D-Y is desirable, there are unique challenges to overcome before this is possible. First, there is a normative question with respect to whose preferences should be elicited. The instrument itself has been designed for self-completion by younger individuals, yet the use of adult general population preferences in health state valuation studies is commonplace (Rowen et al., 2017) with samples typically consisting of individuals aged 18 and above. There is an ongoing debate in the health state valuation literature regarding whose preferences to elicit, with adult general population samples often justified in countries with publicly-funded health care systems on the basis that they represent taxpayers' preferences (Versteegh and Brouwer, 2016). In many countries, younger individuals (under 18) are typically not taxpayers and are often ineligible to vote. It has also been argued that general population preferences are important because all members of the public are potential users of health care services. However, it follows that adults are not potential users of health care services for younger people and therefore the use of adult preferences may not be appropriate on this basis. On the other side of the debate, it has been argued that preferences should ideally be 'informed' based upon experience with the health state that is being valued. Whilst preferences elicited from adult general population samples are not what is meant by 'experienced-based values' (Brazier et al., 2017), and it has been suggested that such samples are not necessarily fully informed (Karimi et al., 2017), adults are arguably better informed about the impact of ill health on health-related quality of life than younger people on average. Nonetheless, given the types of setting in which the EQ-5D-Y would be used, it can still be argued that it would be inappropriate to base resource allocation decisions that affect younger populations on adult preferences alone.

Second, regardless of the stance taken on the normative issue, there are challenging questions with respect to the design of the valuation tasks. If the preferences of children or adolescents are considered relevant, it must be ensured that the valuation task is suitable for them to complete. However, traditional methods that include complex trade-offs involving death, such as time trade-off (TTO) and standard gamble (SG), may be too cognitively challenging for younger individuals. Regardless, even if the decision is made to elicit preferences from a sample of adults, it is not necessarily clear how the valuation task should be framed. For example, adults could be asked to value their own health, described using the EQ-5D-Y instrument, or they could be asked to consider the health of a child when completing the valuation task. It is plausible that the perspective of the task will influence respondents' preferences.

Different valuation exercises to obtain adult preferences for states from child instruments have used these alternative perspectives. For instance, the development of the value set for the CHU-9D instrument asked adults to imagine themselves in the health state for their rest of their lives (Stevens,

2012) whereas a similar exercise for the HUI2 asked adults to imagine that they were a child aged 10 years, and that they would expect to live for another 60 years (McCabe et al., 2005). In the valuation exercise for the EQ-5D-Y in the US, adult respondents were asked to imagine a 7 or 10-year-old child experiencing the health state (Craig et al., 2016). In addition to the differing ages for which the health states are to be considered, the differences in the duration of the states in itself would be expected to exert an influence over the values generated. Two recent studies that explored the potential differences when using different perspectives show contradictory results with respect to which perspective (adult/child) results in higher/lower utility values (Kind et al., 2015; Kreimeier et al., 2018). However, both suggest that different perspectives lead to different conclusions. In addition, Kreimeier et al. (2015) found that using TTO methods in this context was problematic because of an apparent unwillingness of adults to trade life years for a child, leading to relatively high values for poor health states in the child perspective arm.

The alternative approach of asking younger individuals to value health states has also been attempted. In fact, despite the cognitive demands, a recent review about the reliability, validity, and feasibility of direct elicitation of children's preferences for health states identified SG and TTO as the main techniques used to obtain preferences from children (Crump et al., 2017). Although there is evidence that adolescents living with a chronic condition (cystic fibrosis) can complete successfully SG/TTO tasks (Yi et al., 2003), in general this is not the case for young members of the general population. It should also be noted that these elicitation methods engage participants with tasks that involve death, which introduces ethical concerns when presenting them to young individuals (Ratcliffe et al., 2011).

Ultimately, the question of whose preferences to elicit when conducting preference elicitation exercises to generate data for a value set for the EQ-5D-Y is only important if there are differences in the preferences of adults and younger individuals. Two studies have reported that, when presented with the same health state, adults and adolescents exhibited systematically different preferences (Saigal et al., 1999; Wasserman et al., 2005). In both cases, SG was used but no formal assessment of the feasibility of using this technique in adolescents was conducted. Therefore, given the challenges associated with SG and TTO, the use of an alternative method such as discrete choice experiments (DCEs) could be preferable when making such comparisons. DCEs have been gaining popularity for eliciting preferences to EQ-5D states in recent years (Stolk et al., 2010; Krabbe et al., 2014) as they are less cognitively demanding than the alternatives and do not require consideration of duration, or trade-offs with death. However, the standard DCE technique produces relative preferences on a latent, or undefined, scale (which cannot be used in QALY calculations). Further work is therefore required to anchor the results on the standard scale where 0 represents dead and 1 represents 'perfect health'.

Given the gaps in the literature, this study aimed to determine whether preferences differ between adults and adolescents by eliciting preferences from samples of both groups using a DCE.

2 Methodology

2.1 Overview

Two online surveys were administered, one to a sample of adults and another to a sample of adolescents. The surveys were developed in collaboration with epiGenesys, a software development company. Both surveys comprised the following elements (in order): screening questions; information sheet and informed consent; self-reported health using EQ-5D-Y and EQ-VAS; instructions; 16 paired comparison tasks; three debrief questions; and background questions. The background questions differed slightly between samples, and adults were asked some additional debrief questions relating to the framing of the DCE task. Ethics approval to conduct this study was obtained from the Medical Sciences Inter-Divisional Research Ethics Committee (IDREC) at the University of Oxford.

The remainder of section 2 will describe the sample recruitment, the EQ-5D-Y instrument, and the various components of the DCE design.

2.2 Sample Recruitment

Data were collected from a sample of adult members of the UK general public (target sample size: n=1,000) as well as a sample of UK adolescents members of the general public aged between 11 and 17 (target sample size: n=1,000). All adult respondents were members of an online panel managed by a market research agency, Survey Sampling International. The adolescent respondents were the children of adult panel members. Selected panel members that had not been contacted for the adult survey, but had been identified as having children according to the agency's database, were contacted with an invitation for their children to take part.

Quotas, combined with a targeted recruitment strategy, were used to ensure that the sample was representative of the general population in terms of gender, age, social grade (adult sample only) and nation (within the UK; adult sample only). Respondents were awarded 'panel points' (which can be redeemed for cash vouchers and other rewards) following completion of the survey. Prior to launching the survey it was agreed that any respondents completing the entire survey in less than 2.5 minutes would be excluded from the sample on data quality grounds.

2.3 The EQ-5D-Y Instrument

The EQ-5D-Y instrument consists of five dimensions, each with three severity levels, as detailed in Table 1. In contrast to the commonly used EQ-5D-3L instrument, the 'self-care' dimension is labelled as 'looking after myself' and the 'anxiety/depression' dimension is labelled as 'feeling worried, sad or unhappy' as these were deemed to be more easily understood by younger individuals. However, in the interest of brevity, the 'traditional' labels/codes are used throughout this paper. A total of 243 (3^5) health states are possible when using the EQ-5D-Y.

Table 1. EQ-5D-Y Instrument

DIMENSION	LEVELS	CODING
MOBILITY (WALKING ABOUT)	I have no problems walking about	MO1
	I have some problems walking about	MO2
	I have a lot of problems walking about	MO3
LOOKING AFTER MYSELF[1]	I have no problems washing or dressing myself	SC1
	I have some problems washing or dressing myself	SC2
	I have a lot of problems washing or dressing myself	SC3
DOING USUAL ACTIVITIES (FOR EXAMPLE, GOING TO SCHOOL, HOBBIES, SPORTS, PLAYING, DOING THINGS WITH FRIENDS OR FAMILY)	I have no problems doing my usual activities	UA1
	I have some problems doing my usual activities	UA2
	I have a lot of problems doing my usual activities	UA3
HAVING PAIN OR DISCOMFORT	I have no pain or discomfort	PD1
	I have some pain or discomfort	PD2
	I have a lot of pain or discomfort	PD3
FEELING WORRIED, SAD OR UNHAPPY[2]	I am not worried, sad or unhappy	AD1
	I am a bit worried, sad or unhappy	AD2
	I am very worried, sad or unhappy	AD3

[1]Referred to as 'self-care' by convention; [2]Referred to as 'anxiety/depression' by convention.

2.4 The Discrete Choice Experiment

2.4.1 Overview

The DCE required respondents to make a choice between two EQ-5D-Y health states labelled as options A and B. All five dimensions of the EQ-5D-Y were included as attributes, along with the three severity levels for each dimension.

In terms of the framing of the choice tasks, the adult sample were asked:

> *"Considering your views about a 10-year-old child: which do your prefer, A or B?"*

In contrast, the adolescent sample were asked:

> *"Which do you prefer, A or B?"*

The visual presentation of the choice tasks were designed to mimic the format used for DCE tasks in the EuroQol Group's international EQ-5D-5L valuation protocol, often referred to as the EQ-VT protocol (Oppe et al., 2014).

2.4.2 Experimental Design

The experimental design took the form of a Bayesian efficient design, with main effects and all two-way interactions, minimal number of unrealistic health states, overlapping of health states in two dimension levels, and good level and utility balance. Each respondent completed 15 tasks and the design was divided in 10 blocks yielding a total of 150 pairs. The design allowed the estimation of a multinomial logit model with 50 parameters (10 main effects parameters and all 40 two-way interaction parameters).

A two-step procedure was undertaken to select the 150 pairs included in the final design. In the initial step, we randomly selected 150 pairs that maximised the Fisher information matrix and asked 127 adult respondents to complete these pairs in a soft launch. We did not use priors or a Bayesian framework for the initial step as the lack of appropriate prior information could bias the initial design (Hensher et al., 2015). A multinomial logit model estimating main effects and all two-way interactions was used to model the data from the soft launch. The resulting coefficients were used as the basis for priors to inform the design for the main launch (second step in the experimental design). For this final design, we used two set of priors, one assuming a main effects only model (10 parameters) and another using main effects plus all two-way interactions (10 + 40 parameters). We simulated 1,000 designs, calculated the mean error for both sets of priors and identified the five designs with the smallest mean errors for the main effects model. From these five, we selected the final design identifying the one with the smallest error in the main effect plus interactions model. Blocking for the 150 pairs in both steps was implemented by minimising the variance of the level balance between blocks. The design is available on request to the authors.

2.4.3 Discrete Choice Modelling

Choice data are typically modelled using a random utility model (RUM) framework, where the utility obtained by decision-maker n choosing alternative j is given by Equation 1.

$$U_{nj} = V_{nj} + \varepsilon_{nj} \qquad (1)$$

Where Vnj is an observable (or deterministic) component made up of the attributes of the alternatives X_{nj} and observable characteristics of the decision-maker Z_n. ε_{nj} is an unknown (or stochastic) component and treated as random. It follows that the probability that the decision-maker n chooses alternative i is:

$$P_{ni} = Pr(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj}) \ \forall j \neq i \qquad (2)$$

Different choice models are obtained from different assumptions about the distribution of the random terms. The most commonly used choice model, the multinomial logit (MNL), assumes that the random terms are independent and identically distributed (IID) type one extreme value and suffers from the restrictive independence of irrelevant alternatives (IIA) property. Additionally, the MNL model assumes that preferences are homogenous across individuals, unless systematic differences across participants are included in the observable component of utility (e.g. gender, age). Due to this, alternative models are typically preferred on the basis that actual choice behaviour can be better represented by flexible models that attempt to control for various sources of random heterogeneity.

There are two types of random heterogeneity in particular that alternative choice models typically try to account for. The first, preference heterogeneity, occurs if individuals' preferences differ from one another for reasons beyond differences in observable characteristics. The second, scale heterogeneity, is a specific type of correlation across utility coefficients. It occurs when the impact of factors not included (in the model) affect individuals differently, giving the impression that some individuals' responses are 'more random' than others (Hess and Train, 2017). Several suggestions exist in the econometric literature for incorporating preference and scale heterogeneity in a discrete choice model (McFadden and Train, 2000; Fiebig et al., 2010; Zhou et al., 2017). Previous research in the adult and adolescent samples used in this study have shown that mixed logit (MIXL), generalised multinomial logit and latent class models were associated with better fit and prediction accuracy than MNL (Rivero-Arias et al., 2017; forthcoming). The study also showed that there were negligible benefits of one model over another in terms of deriving a latent scale value set. In this study we selected the MIXL model as the basis of our comparison between adult and adolescent samples. MIXL models allow parameters to be estimated for each respondent in the sample and hence take preference heterogeneity into account (McFadden and Train, 2000). MIXL models can also take scale heterogeneity into account if a full covariance matrix is estimated.

A linear, additive utility function was estimated with all variables dummy coded and 'level 1s' used as base levels.

2.4.4 Preference Comparisons

An increasingly well-documented issue when comparing the preferences of different samples using DCE data is the confounding between preference and scale (Vass et al., 2018; Wright et al., 2018). The implication of this is that it becomes difficult to determine whether different coefficients occur as a result of differences in preferences, or as a result of differences in scale.[1] It is possible to determine whether differences in scale exist between samples using the Swait-Louviere test; this is typically conducted using MNL models and is the approach used in this study (Vass et al., 2018). However, this does not allow for both scale and preference heterogeneity to be controlled for. In fact, it has been argued that it is impossible to disentangle the two (Hess and Train, 2017).

Therefore, in order to control for scale and preference heterogeneity as best as possible whilst proceeding to make comparisons between the adult and adolescent samples, two approaches will be used. The first approach is to examine the relative importance of the different dimensions. This approach involves estimating the latent utility range for each attribute, and dividing this by the overall latent utility range (for all attributes). Pooled models are used to test for statistically significant differences in the relative importance figures between the two samples. The second approach is to use the MIXL model coefficients to generate an implied ranking for each attribute level for the two samples, and to compare these rankings.

2.4.5 Logic Checks

In addition to the choice sets drawn from the experimental design, all respondents completed one further 'fixed pair' in which one health state (11122) could be considered to logically dominate the other (22233). Given concerns about the quality of internet survey data (Rowen et al., 2016), it can be useful to include a task that can act as a 'rationality check', helping to identify respondents whose choices suggest a poor level of attentiveness, engagement or understanding. A large proportion of respondents failing to choose the dominant option could be considered to be a sign of poor data quality. The 'fixed' pair data were excluded from the modelling exercise. Given that the two alternatives in each task were EQ-5D-Y health states, a crude data quality check would be to examine the proportion of respondents that chose health states with a lower 'level sum score' (LSS). The LSS for any given health state is calculated by taking the sum of its levels. For example, the LSS for 11111 is 5 and the LSS for 33333 is 15. It follows that a higher LSS corresponds to a more severe state. The larger the difference in LSS between any two health states, the greater the expectation that a respondent would choose the option with the lower LSS. However, it should be noted that this is not as clear-cut a comparison as the dominance test, as respondents may have reason to believe that a health state with a higher LSS is preferable to another with a lower LSS in some circumstances.

1 Note: The MIXL model only controls for scale heterogeneity within samples and does not control for scale heterogeneity such that comparisons between samples can be made.

3 Results

3.1 Response Rates and Sample Composition

Main data collection for the adult sample was carried out in February/March 2017. Of the 1,187 individuals who accessed the survey, 87 (7.3%) declined consent, 72 (6.1%) started but did not provide a complete set of data, and 28 (2.4%) completed the survey in less than the agreed minimum time of 2.5 minutes. This left a total of 1,000 respondents for analysis. Main data collection for the adolescent sample was carried out in November/December 2017. Of the 1,449 individuals who accessed the survey, 192 (13.2%) were outside the eligible age range, 136 (9.4%) declined consent, 56 (3.9%) started but did not provide a complete set of data, and 60 (4.1%) completed the survey in less than the agreed minimum time. This left a total of 1,005 respondents for analysis.

Background characteristics of the two samples are summarised in Table 2. Quotas were used to generate representative samples of the UK general population. By construction, the adult sample was representative of the UK general population in terms of age group, gender, social grade and nation, and the adolescent sample was representative of the UK adolescent population in terms of age group (i.e. split between 11-to-14 year olds and 15-to-17 year olds) and gender. After excluding speeders, the mean (median) amount of time taken to complete the survey was 11 minutes (7 minutes) for the adult sample, and 9 minutes (6 minutes) for the adolescent sample.

The key difference between the two samples can be seen when looking at the self-reported EQ-5D-Y. 58% of the adolescent sample reported the best health state (11111) in contrast to 15% of the adult sample. The differences can be seen more clearly in Figure 1, which illustrates the percentage of respondents from each sample that reported each level on each dimension. As expected, a far greater proportion of the adult sample reported levels 2 or 3 in the various dimensions compared with the adolescent sample.

Table 2. Sample Background Characteristics

		ADULT SAMPLE N=1,000 (%)	ADOLESCENT SAMPLE N=1,005 (%)	GENERAL POPULATION
GENDER	Female	512 (51.2%)	494 (49.1%)	51%
	Male	488 (48.8%)	511 (50.9%)	49%
AGE	11		78 (7.8%)	15%
	12		132 (13.1%)	14%
	13		181 (18.0%)	14%
	14	N/A	174 (17.3%)	14%
	15		162 (16.1%)	14%
	16		139 (13.8%)	14%
	17		139 (13.8%)	15%
	18-29	199 (19.9%)		20%
	30-44	272 (27.2%)	N/A	25%
	45-59	255 (25.5%)		26%
	60+	274 (27.4%)		30%
NATION	England	845 (84.5%)	857 (85.3%)	84%
	Scotland	85 (8.5%)	72 (7.2%)	
	Wales	49 (4.9%)	58 (5.8%)	16%
	Northern Ireland	21 (2.1%)	18 (1.8%)	
SOCIAL GRADE	Higher (ABC1)	542 (54.2%)		55%
	Lower (C2DE)	458 (45.8%)		44%
FAMILY AFFLUENCE SCALE	Low Score (0-2)		30 (3%)	N/A
	Medium Score (3-5)		456 (45%)	N/A
	High Score (6-9)		519 (52%)	N/A
SELF-REPORTED HEALTH (EQ-5D-Y)	Health State 11111	148 (14.8%)	587 (58.4%)	N/A
	All Other Health States	852 (85.2%)	418 (41.6%)	N/A

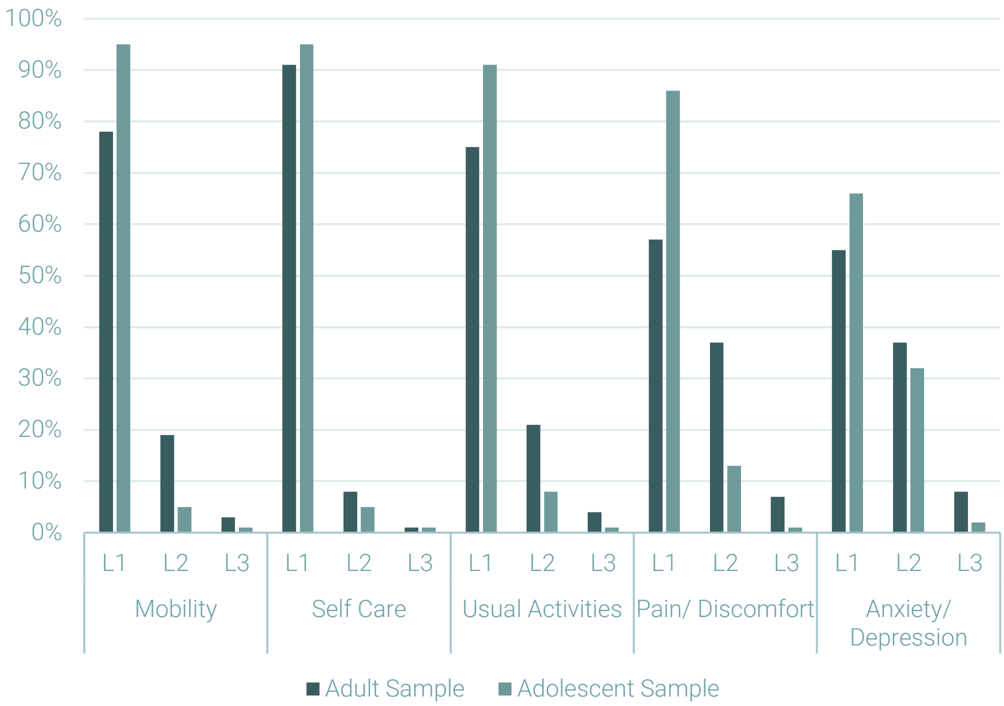
Higher (ABC1) indicates that the Chief Income Earner in the respondent's household works in a managerial, administrative or professional occupational group; Lower (C2DE) indicates that they are a skilled, semi-skilled or unskilled manual worker, stated pensioner, casual/lowest grade worker, or unemployed with state benefits only.

General population gender stats refer to % of the entire UK population, whereas age stats refer to % of the 11-to-17-year-old and 18+ year populations, respectively

General population gender and age stats taken from:
Office for National Statistics, 2017. Population estimates for UK, England and Wales, Scotland and Northern Ireland. [dataset] Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland [accessed 4 Oct 2017]

General population social grade stats taken from:
National Readership Survey, 2016. Social Grade. Available at: http://www.nrs.co.uk/nrs-print/lifestyle-and-classification-data/social-grade/ [Accessed 4 May 2018]

Figure 1. Self-Reported EQ-5D-Y



OFFICE OF HEALTH ECONOMICS RESEARCH

3.2 Discrete Choice Experiment Results

The main regression results can be found in Table 3. In both samples, the coefficient for every dimension included in the MNL and MIXL models were negative and statistically significant at the 1% level. In addition, in the MIXL models for both samples, every standard deviation except MO2 was statistically significant at the 1% level indicating evidence of preference heterogeneity, spanning most dimensions and levels of the EQ-5D-Y. Such a result is an indication of the suitability of the MIXL model over the MNL model in this study, which is also captured by better fit of the data with lower log-likelihood, AIC and BIC. When implementing the Swait-Louviere test using the MNL model, it was found that differences in coefficients were not explained solely by differences in scale, indicating that preferences do exist between the two samples.

Using the MIXL model results, implied rankings for the different levels (reflecting the sizes of the coefficients) can be inferred; the implied rankings can be found in Table 4. In both samples the two most important attribute levels were PD3 and AD3. Similarly, the three least important attribute levels were UA2, MO2 and SC2. Despite this consistency at the extreme ends of the spectrum, the rankings of the remaining five attribute levels differed entirely. For example, MO3 is considered more important than UA3 for adolescents whereas the opposite is the case for adults. Therefore, preferences do appear to differ somewhat between the two samples when considering the levels of the dimensions. However, the overall relative importance of the five dimensions is not too dissimilar, as illustrated by Figure 2. In both samples, pain/discomfort is the most important attribute followed by anxiety/depression and the least important attribute is self-care. There is a minor difference in the implied ranking between the other two attributes; for adults usual activities is slightly more important than mobility, whereas the opposite is the case for adolescents.

Figure 3 shows a scatter plot of observed versus predicted choice probabilities for each of the 150 DCE pairs. When comparing the two samples, it can be seen that there are fewer observations in the middle range (where predicted and observed probabilities are around 0.5) for the adult sample when compared with the adolescent sample. This suggests that, in the adult sample, there were fewer cases where the probability of respondents choosing each option was relatively equal, compared with the adolescent sample.

A comparison of responses to the pair containing the dominance test showed little difference between the two samples. Approximately 90% of the adult sample chose the dominant option relative to 88% in the adolescent sample. Figure 4 illustrates that, when comparing the responses based on differences in LSS, the expected pattern was observed for both samples. Adolescents were generally slightly less likely to choose the 'more severe' option in the tasks, and this was exacerbated when the difference in LSS was very small. For example, when option A had a LSS that was one point lower than that of option B (i.e. option A was 'less severe' than option B), 76% of adults chose option A relative to 68% of adolescents.

Responses to the debrief questions suggested that an equal proportion (51%) of both the adult and adolescent samples disagreed or strongly disagreed that the tasks were difficult. However, 67% of the adult sample disagreed or strongly disagreed that they found it difficult to tell the difference between the health descriptions, in contrast to 55% of the adolescent sample. Finally, 52% of the adult sample disagreed or strongly disagreed that they found it difficult to imagine the health problems described, in contrast to 35% of the adolescent sample.

Table 3. Regression Results

	ADULT SAMPLE			ADOLESCENT SAMPLE		
	MNL	MIXL		MNL	MIXL	
	Coeff.	*Coeff.*	*Std. Dev.*	*Coeff.*	*Coeff.*	*Std. Dev.*
MO2	-0.158***	-0.408***	0.547***	-0.255***	-0.407***	0.086
	(0.048)	(0.067)	(0.112)	(0.046)	(0.062)	(0.151)
MO3	-0.611***	-1.200***	1.246***	-0.896***	-1.419***	1.166***
	(0.077)	(0.114)	(0.158)	(0.074)	(0.106)	(0.186)
SC2	-0.247***	-0.365***	0.240***	-0.196***	-0.332***	0.481***
	(0.039)	(0.057)	(0.083)	(0.037)	(0.053)	(0.075)
SC3	-0.592***	-0.979***	0.806***	-0.723***	-1.123***	1.148***
	(0.065)	(0.090)	(0.123)	(0.063)	(0.090)	(0.119)
UA2	-0.372***	-0.607***	0.702***	-0.310***	-0.496***	0.615***
	(0.042)	(0.061)	(0.082)	(0.040)	(0.054)	(0.077)
UA3	-0.894***	-1.478***	1.171***	-0.819***	-1.328***	1.326***
	(0.051)	(0.090)	(0.097)	(0.051)	(0.085)	(0.121)
PD2	-0.581***	-1.128***	1.100***	-0.492***	-0.818***	0.865***
	(0.043)	(0.077)	(0.080)	(0.039)	(0.060)	(0.081)
PD3	-1.553***	-3.057***	2.560***	-1.414***	-2.319***	1.996***
	(0.075)	(0.159)	(0.138)	(0.064)	(0.114)	(0.140)
AD2	-0.602***	-0.951***	0.900***	-0.363***	-0.566***	0.722***
	(0.043)	(0.070)	(0.095)	(0.039)	(0.056)	(0.087)
AD3	-1.504***	-2.592***	2.048***	-1.310***	-2.162***	1.952***
	(0.069)	(0.131)	(0.121)	(0.065)	(0.114)	(0.138)
N	30,000	30,000		30,150	30,150	
LL	-8,300	-7,225		-8,907	-8,013	
BIC	16,703	15,120		17,917	16,697	

Standard errors in parentheses: ***p<0.01; **p<0.05, *p<0.1; MNL: multinomial logit; MIXL: mixed logit; Coeff.: coefficient; Std. Dev.: standard deviation; MO2: I have some problems walking about; MO3: I have a lot of problems walking about; SC2: I have some problems washing or dressing myself; SC3: I have a lot of problems washing or dressing myself; UA2: I have some problems doing my usual activities; UA3: I have a lot of problems doing my usual activities; PD2: I have some pain or discomfort; PD3: I have a lot of pain or discomfort; AD2: I am a bit worried, sad or unhappy; AD3: I am very worried, sad or unhappy; LL: log-likelihood; AIC: Akaike information criteria; BIC: Bayesian information criteria.

Table 4. Implied Rankings from the Mixed Logit Models

	ADULT SAMPLE	ADOLESCENT SAMPLE
Most Important	PD3	PD3
↑	AD3	AD3
	UA3	MO3
	MO3	UA3
	PD2	SC3
	AD2	PD2
	SC3	AD2
	UA2	UA2
↓	MO2	MO2
Least Important	SC2	SC2

Figure 2. Relative Importance of the Dimensions

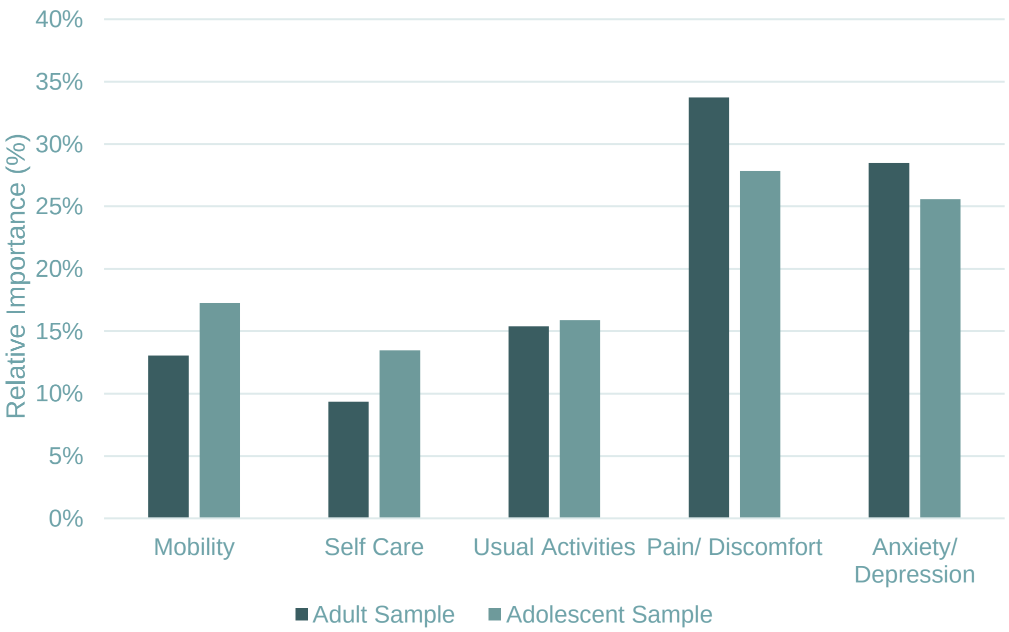
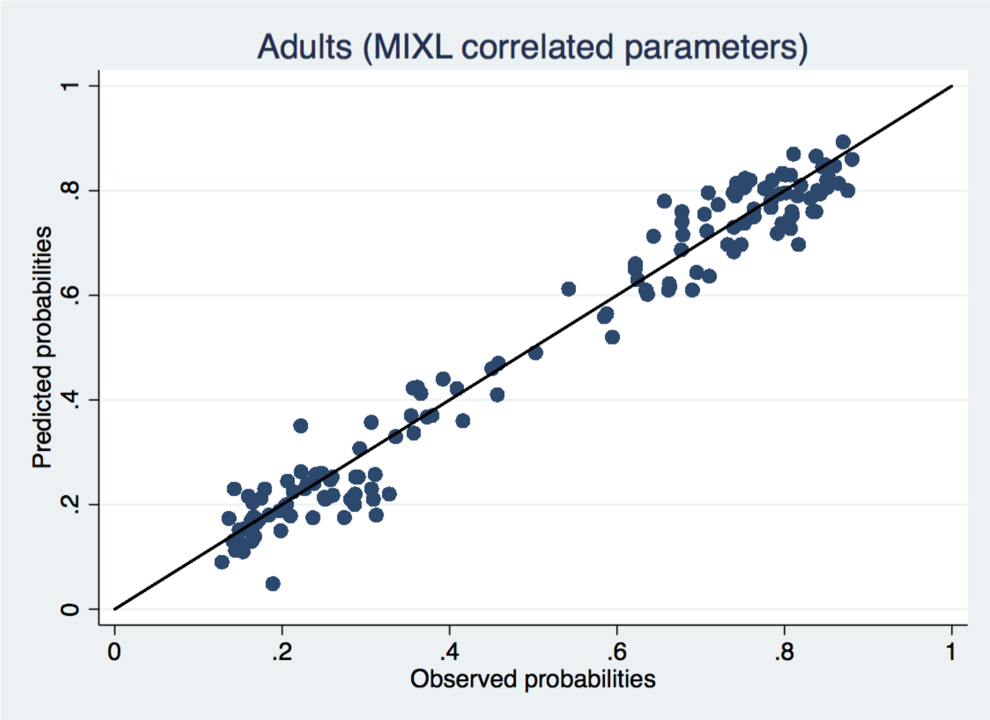


Figure 3. Predicted Probabilities



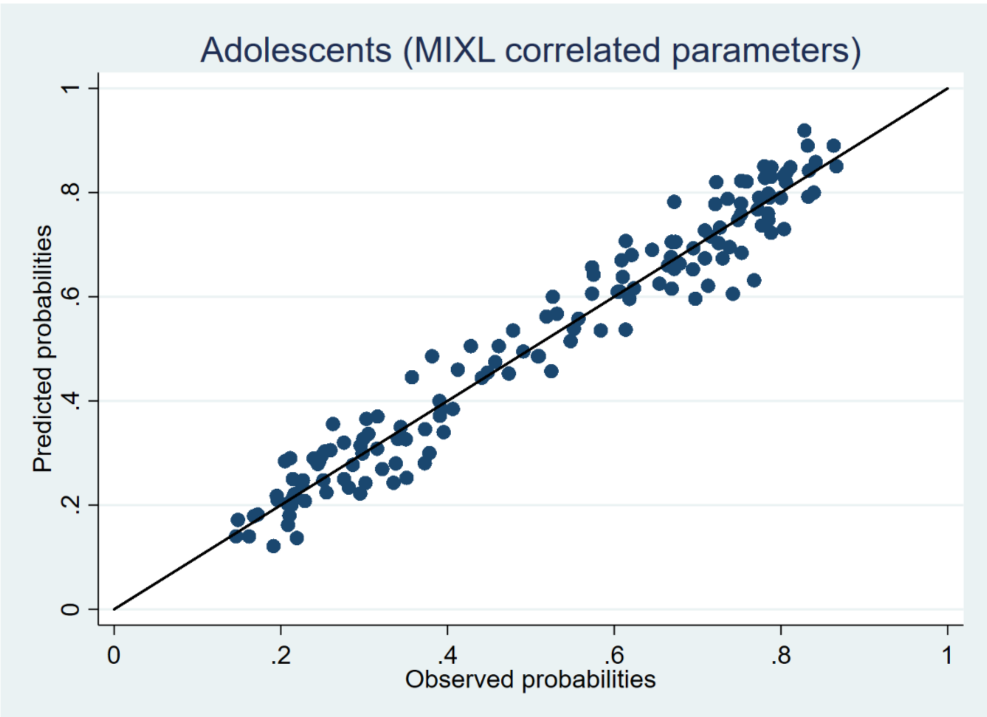
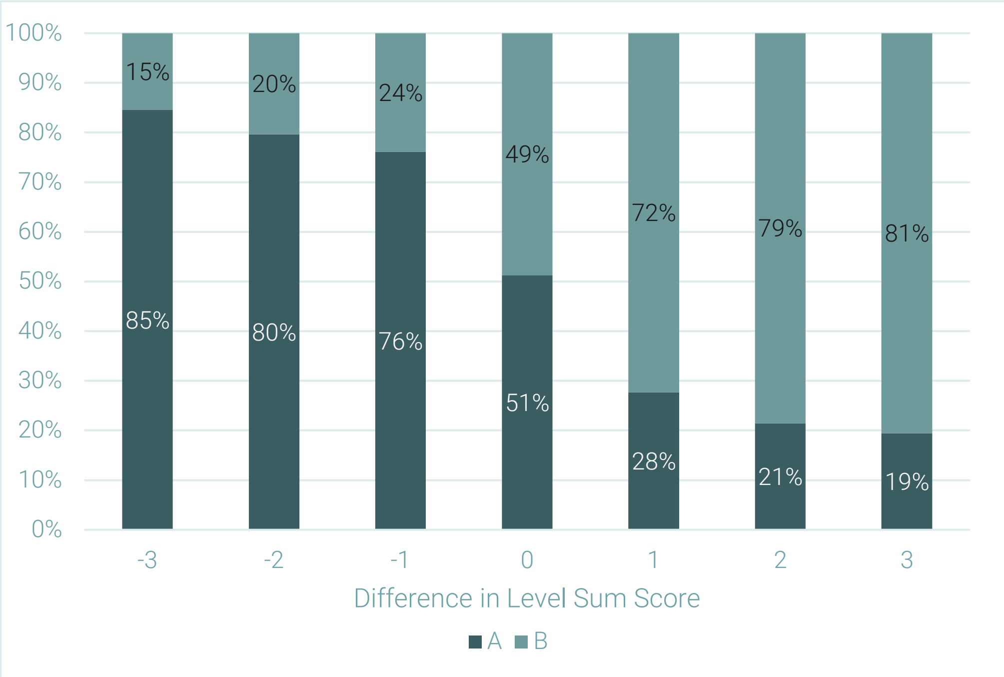
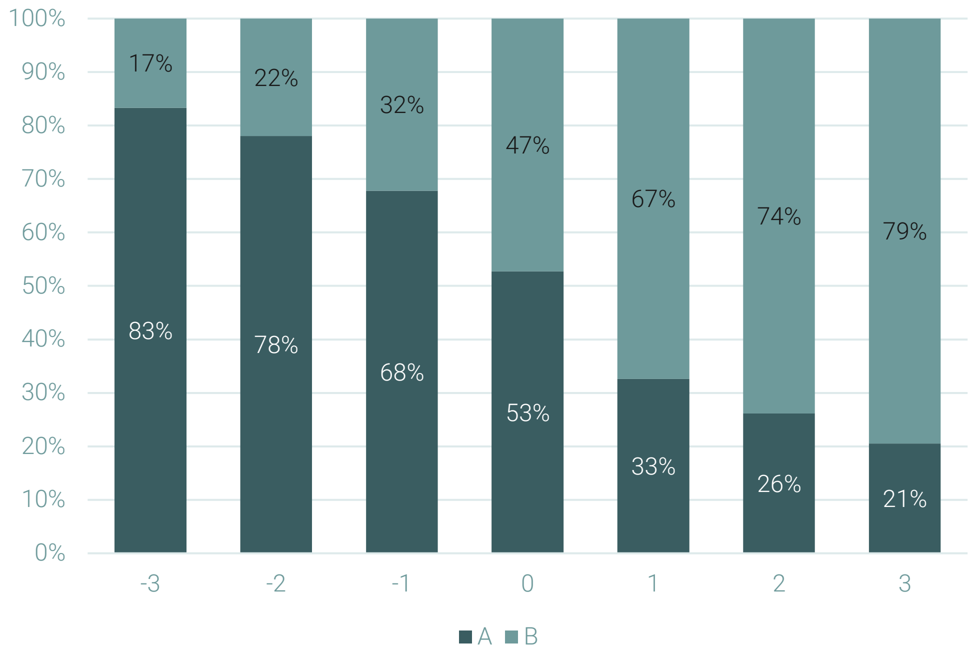


Figure 4. Proportion of Respondents Choosing A/B

Adults



Adolescents



OFFICE OF HEALTH ECONOMICS
RESEARCH

4 Discussion

In this study, we have implemented MNL and MIXL models to estimate latent scale utilities for EQ-5D-Y states based on stated preference data obtained from a DCE that was provided to both a sample of adults and a sample of adolescents. The majority of analyses indicate that preferences differ slightly between the two samples. In particular, it appears that adolescents are more concerned with mobility issues than those in the adult sample. This may be due to differences in experience; fewer adolescents are likely to have suffered with mobility issues relative to adults. It could be the case that adults are less concerned about mobility issues on average, which could be due to having different mobility expectations (possibly due to adaptation).

It is interesting to observe that there were fewer choice probabilities centred on 0.5 for adults relative to adolescents. This might imply that the adolescent sample on average are less consistent than the adult sample. For example, for any given choice, a larger majority of the adult sample may have chosen a particular alternative relative to the adolescent sample (i.e. a smaller majority). This fits relatively well with the finding that adolescents typically were less likely to choose the option with a lower LSS when the difference in LSS between the two options was small. These findings suggest that adolescents may have struggled with the task to a greater extent than adults, despite an equal proportion of the two samples reporting that they did not find the tasks difficult. It could be the case that this is better explained by the relatively greater difficulty that adolescents had compared with adults when it came to telling the difference between the health descriptions and imagining the health problems that were described. Regardless, it is important to note that these are minor differences and that the proportion of adolescents passing the dominance test was almost identical to the proportion of adults passing. Ultimately, it would seem fair to conclude that the adolescent sample did not struggle significantly with the DCE, and that this methodology would appear to be suitable for use in this age range.

Given that it appears to be feasible to elicit preferences from adolescents, a logical question follows relating to how this information could, or should, be used. One suggestion might be to exclusively use EQ-5D-Y value sets based on adolescent preferences for evaluations of interventions aimed at younger populations. However, an alternative might be to combine the data to create a single general population EQ-5D-Y value set, which includes preference data from a smaller (representative) number of individuals aged between 11 and 17 years.

This study has a number of limitations. First, the perspective of the task differed between the two samples. Adults were asked to express preferences with respect to another individual, whereas adolescents were asked for their own individual preferences. Various theoretical frameworks have highlighted the importance of differences in perspective when eliciting preferences in health (Dolan et al., 2003; Tsuchiya and Watson, 2017; Cubi-Molla et al., 2018). It is therefore important to highlight that *both* the respondent sample *and* the perspective of the task differed in our study, which reduces our ability to accurately determine why preferences between the two samples differ. Another limitation is that the adult sample were asked to think about a 10-year-old child experiencing the health states to be valued, without specifying who that child is. Our intention was to avoid specific ways of framing the questions that may have limited the generalisability of the preferences elicited. However, the risk with this approach is that we do not know the cognitive process employed by respondents in completing the tasks – for example, some may have considered themselves as a 10-year-old, or considered a 10-year-old they know, or imagined a hypothetical 10-year-old. The approaches might differ across respondents, and could have been different had the 'reference child' been framed in a different manner. For example, the age of the 'reference child' may have made a difference. Nobody in the adolescent sample was aged under 11 and therefore there was no direct comparison. Adults may have expressed different preferences had the 'reference child' been older.

Ultimately it is not possible to tell the extent to which the framing of the question may have played a role in the differences that were observed between the two samples, and only further research will be able to uncover the influence that framing may have. In addition, another limitation is that the DCE tasks did not include any consideration of dead or the duration of the health states, and the latent scale results reported here are therefore not anchored in a manner that would enable them to be used in the estimation of QALYs, without further information. This means that comparisons are limited to the relative importance of the different levels rather than comparisons of (anchored) utilities, which would be more meaningful. However, the inclusion of dead or duration may have made the task too difficult for the adolescent sample and may have raised ethical issues. Post-hoc anchoring of latent scale values may enable a value set to be created at a later date.

5 Conclusion

Our evidence suggests that adolescents' preferences differ from those of adults taking the perspective of a child. It may be that these differences exist due to the relative experience of adults, who might have a better understanding of ill health and its effect on health-related quality of life. However, a normative argument can be made that adolescents' preferences should be considered in decision-making that is directly relevant to them. Whilst the cognitive demands of other valuation methods may have ruled this possibility out, this study provides evidence to suggest that adolescents are capable of completing a DCE. Future research should explore further the possible differences that may occur in value sets as a result of these latent scale differences.

References

Brazier, J., Rowen, D., Karimi, M., Peasgood, T., Tsuchiya, A. and Ratcliffe, J., 2017. Experience-based utility and own health state valuation for a health state classification system: why and how to do it. *The European Journal of Health Economics*, 19(6), pp.881-891.

Craig, B.M., Greiner, W., Brown, D.S., and Reeve, B.B., 2016. Valuation of Child Health-Related Quality of Life in the United States. *Health Economics*, 25(6), pp.768-77.

Crump, R.T., Beverung, L.M., Lau, R., Sieracki, R. and Nicholson, M., 2017. Reliability, Validity, and Feasibility of Direct Elicitation of Children's Preferences for Health States. *Medical Decision Making*, 37(3), pp.314–326.

Cubi-Molla, P., Shah, K. and Burström, K., 2018. Experience-Based Values: A Framework for Classifying Different Types of Experience in Health Valuation Research. *The Patient*, 11(3), pp.253–270.

Dolan, P., Olsen, J.A., Menzel, P. and Richardson, J., 2003. An inquiry into the different perspectives that can be used when eliciting preferences in health. *Health Economics*, 12(7), pp.545–551.

Fiebig, D. G., Keane, M. P., Louviere, J. and Wasi, N., 2010. The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity. *Marketing Science*, 29(3), pp.393-421.

Hensher, D.A., Rose, J.M., and Greene, W.H., 2015. *Applied Choice Analysis*. 2nd edition ed. Cambridge: Cambridge University Press.

Hess, S. and Train, K., 2017. Correlation and scale in mixed logit models. *Journal of Choice Modelling*, 23(Supplement C), pp.1–8.

Karimi, M., Brazier, J. and Paisley, S., 2017. Are preferences over health states informed? *Health and Quality of Life Outcomes*, 15(1).

Kind, P., Klose, K., Gusi, N., Olivares, P.R., and Greiner, W., 2015. Can adult weights be used to value child health states? Testing the influence of perspective in valuing EQ-5D-Y. *Quality of Life Research*, 24(10), pp.2519-2539.

Krabbe, P.F., Devlin, N.J., Stolk, E.A., Shah, K.K., Oppe, M., van Hout, B., Quik, E.H., Pickard, A.S., and Xie, F., 2014. Multinational evidence of the applicability and robustness of discrete choice modeling for deriving EQ-5D-5L health-state values. *Medical Care*, 52(11), pp.935-943.

Kreimeier, S., Oppe, M., Ramos-Goñi, J.M., Cole, A., Devlin, N., Herdman, M., Mulhern, B., Shah, K.K., Stolk, E., Rivero-Arias, O. and Greiner, W., 2018. Valuation of EuroQol five-dimensional questionnaire, youth version (EQ-5D-Y) and EuroQol five-dimensional questionnaire, three-level version (EQ-5D-3L) health states: the impact of wording and perspective. *Value in Health*, 21(11), pp.1291-1298.

McCabe, C., Stevens, K., Roberts, J., and Brazier, J., 2005. Health state values for the HUI 2 descriptive system: results from a UK survey. *Health Economics*, 14(3), pp.231-44.

McFadden, D. and Train, K., 2000. Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15(5), pp.447-470.

Oppe, M., Devlin, N.J., van Hout, B., Krabbe, P.F.M. and de Charro, F., 2014. A program of methodological research to arrive at the new international EQ-5D-5L valuation protocol. *Value in Health*, 17(4), pp.445–453.

Ratcliffe, J., Couzner, L., Flynn, T., Sawyer, M., Stevens, K., Brazier, J. and Burgess, L., 2011. Valuing Child Health Utility 9D health states with a young adolescent sample: a feasibility study to compare best-worst scaling discrete-choice experiment, standard gamble and time trade-off methods. *Applied Health Economics and Health Policy*, 9(1), pp.15-27.

Ravens-Sieberer, U., Wille, N., Badia, X., Bonsel, G., Burström, K., Cavrini, G., Devlin, N., Egmar, A.-C., Gusi, N., Herdman, M. et al., 2010. Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study. *Quality of Life Research*, 19(6), pp.887–897.

Rivero-Arias, O., Shah, K.K., Ramos-Goñi, J.M., Mott, D.J., Devlin, N.J., 2017. *Estimating latent scale discrete choice utilities to develop an EQ-5D-Y value set in the UK.* Paper presented at the 34th EuroQol Plenary Meeting 2017, Barcelona.

Rivero-Arias, O., Mott, D.J., Shah, K.K., Ramos-Goñi, J.M., Devlin, N.J., forthcoming. *Comparing Alternative Choice Models of Heterogeneity in the Valuation of Children's Health.* OHE Research Paper. London: OHE.

Rowen, D., Brazier, J., Mukuria, C., Keetharuth, A., Risa Hole, A., Tsuchiya, A., Whyte, S., and Shackley, P., 2016. Eliciting societal preferences for weighting QALYs for burden of illness and end of life. *Medical Decision Making*, 36(2), pp.210-222.

Rowen, D., Azzabi Zouraq, I., Chevrou-Severac, H. and van Hout, B., 2017. International Regulations and Recommendations for Utility Data for Health Technology Assessment. *PharmacoEconomics*, 35(Suppl 1), pp.11–19.

Saigal, S., Stoskopf, B.L., Feeny, D., Furlong, W., Burrows, E., Rosenbaum, P.L. and Hoult, L., 1999. Differences in preferences for neonatal outcomes among health care professionals, parents, and adolescents. *JAMA*, 281(21), pp.1991–1997.

Stevens, K., 2012. Valuation of the Child Health Utility 9D Index. *PharmacoEconomics*. 30(8), pp.729-747.

Stolk, E.A., Oppe, M., Scalone, L., and Krabbe, P.F., 2010. Discrete choice modeling for the quantification of health states: the case of the EQ-5D. *Value in Health*, 13(8), pp.1005-1013.

Tsuchiya, A. and Watson, V., 2017. Re-Thinking 'The Different Perspectives That can be Used When Eliciting Preferences in Health'. *Health Economics*, 26(12), e103–e107.

Vass, C.M., Wright, S., Burton, M. and Payne, K., 2018. Scale Heterogeneity in Healthcare Discrete Choice Experiments: A Primer. *The Patient*, 11(2), pp.167–173.

Versteegh, M.M. and Brouwer, W.B.F., 2016. Patient and general public preferences for health states: A call to reconsider current guidelines. *Social Science and Medicine*, pp.66–74.

Wasserman, J., Aday, L.A., Begley, C.E., Ahn, C. and Lairson, D.R., 2005. Measuring health state preferences for hemophilia: development of a disease-specific utility instrument. *Haemophilia*, 11(1), pp.49–57.

Wille, N., Badia, X., Bonsel, G., Burström, K., Cavrini, G., Devlin, N., Egmar, A.-C., Greiner, W., Gusi, N., Herdman, M. et al., 2010. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Quality of Life Research*, 19(6), pp.875–886.

Wright, S.J., Vass, C.M., Sim, G., Burton, M., Fiebig, D.G. and Payne, K., 2018. Accounting for Scale Heterogeneity in Healthcare-Related Discrete Choice Experiments when Comparing Stated Preferences: A Systematic Review. *The Patient*, 11(5), pp.475-488.

Yi, M.S., Britto, M.T., Wilmott, R.W., Kotagal, U.R., Eckman, M.H., Nielson, D.W., Kociela, V.L. and Tsevat, J., 2003. Health values of adolescents with cystic fibrosis. *The Journal of Pediatrics*, 142(2), pp.133-140.

Zhou, M., Thayer, W. M. and Bridges, J.F., 2017. Using Latent Class Analysis to Model Preference Heterogeneity in Health: A Systematic Review. *Value in Health*, 20(5), A339-A339.

About us

Founded in 1962 by the Association of the British Pharmaceutical Society, the Office of Health Economics (OHE) is not only the world's oldest health economics research group, but also one of the most prestigious and influential.

OHE provides market-leading insights and in-depth analyses into health economics & health policy. Our pioneering work informs health care and pharmaceutical decision-making across the globe, enabling clients to think differently and to find alternative solutions to the industry's most complex problems.

Our mission is to guide and inform the healthcare industry through today's era of unprecedented change and evolution. We are dedicated to helping policy makers and the pharmaceutical industry make better decisions that ultimately benefit patients, the industry and society as a whole.

OHE. For better healthcare decisions.

Areas of expertise

- Evaluation of health care policy
- The economics of health care systems
- Health technology assessment (HTA) methodology and approaches
- HTA's impact on decision making, health care spending and the delivery of care
- Pricing and reimbursement for biologics and pharmaceuticals, including value-based pricing, risk sharing and biosimilars market competition
- The costs of treating, or failing to treat, specific diseases and conditions
- Drivers of, and incentives for, the uptake of pharmaceuticals and prescription medicines
- Competition and incentives for improving the quality and efficiency of health care
- Incentives, disincentives, regulation and the costs of R&D for pharmaceuticals and innovation in medicine
- Capturing preferences using patient-reported outcomes measures (PROMs) and time trade-off (TTO) methodology
- Roles of the private and charity sectors in health care and research
- Health and health care statistics

ohe.org

The Office of Health Economics A Company Limited by Guarantee of Registered No. 09848965
OHE Consulting Ltd Registered Company No. 09853113
OHE is a Charity Registration No. 1170829
Registration Office Southside, 105 Victoria Street, London SW1E 6QT