



Self vs. other, child vs. adult. An experimental comparison of valuation perspectives for valuation of EQ-5D-Y-3L health states

S. A. Lipman¹ · V. T. Reckers-Droog¹ · M. Karimi^{1,2} · M. Jakubczyk³ · A. E. Attema¹

Received: 7 March 2021 / Accepted: 31 August 2021 / Published online: 6 October 2021
© The Author(s) 2021

Abstract

Objectives EQ-5D-Y-3L health states are valued by adults taking the perspective of a 10-year-old child. Compared to valuation of adult EQ-5D instruments, this entails two changes to the perspective: (i) child health states are valued instead of adult health states and: (ii) health states are valued for someone else instead of for oneself. Although earlier work has shown that these combined changes yield different values for child and adult health states that are otherwise equal, it currently remains unclear why. Hence, we aimed to disentangle the effects of both changes.

Methods A sample of 205 students (mean age: 19.48) was surveyed. Each respondent completed visual analogue scale (VAS) and time trade-off (TTO) tasks for five EQ-5D-Y-3L states, using four randomly ordered perspectives: (i) self-adult (themselves), (ii) other-adult (someone their age), (iii) self-child (themselves as a 10-year-old), (iv) other-child (a child of 10 years old). We compared how each perspective impacted outcomes, precision and quality of EQ-5D-Y-3L valuation.

Results Overall, differences between perspectives were consistent, with their direction being dependent on the health states and respondents. For VAS, the effect on outcomes of valuation depended on severity, but variance was higher in valuation with child perspectives. For TTO, we observed that EQ-5D-Y-3L states valued on behalf of others (i.e., children or adults) received higher valuations, but lower variances.

Conclusion The use of a different perspective appears to yield systematic differences in EQ-5D-Y-3L valuation, with considerable heterogeneity between health states and respondents. This may explain mixed findings in earlier work.

Keywords Health state valuation · Perspective · EQ-5D-Y · Time trade-off · Child health

JEL Classification I10

Introduction

The growing interest in the use of cost-utility analyses for paediatric populations [22] necessitates measurement and valuation of health-related quality of life (HRQOL) in children and adolescents. Valuation of child and adolescent health, however, is associated with several methodological challenges, specifically associated with the methods,

sample and perspective used for valuation [20, 26, 39, 46]. Although various measures of HRQOL for adult and paediatric populations exist [7], decision bodies in several countries recommend using EQ-5D instruments for measurement and valuation of HRQOL (e.g., [18, 30, 53]). For assessing HRQOL in children aged 8 to 15 years, the EQ-5D-Y-3L is recommended [14] and a protocol to obtain value sets for estimation of quality-adjusted life years (QALYs) in health technology assessment has been released recently [35]. Similar to obtaining value sets for adult EQ-5D instruments (i.e., EQ-5D-3L and -5L), EQ-5D-Y-3L value sets are following protocol obtained using a sample representative of the general adult population. However, whereas for adult instruments respondents value EQ-5D health states for themselves

✉ S. A. Lipman
Lipman@eshpm.eur.nl

¹ Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

² Pharmerit, Rotterdam, The Netherlands

³ SGH Warsaw School of Economics, Decision Analysis and Support Unit, Warsaw, Poland

Table 1 Empirical evidence on the impact of (different) perspectives on EQ-5D-Y-3L health state valuations

Authors	Methods	Perspectives		Sample(s) and design	Differences	Attributable to
		$\Delta(S-O)$	$\Delta(A-C)$			
Kind et al. [19]	VAS	OA SA	OC	Adults: SA and OC or SA and OA	OA > SA > OC	$\Delta(S-O)$ and $\Delta(A-C)$ or order effects
Kreimeier et al. [21]	TTO and DCE	$\Delta(S-O)$	OC	Adults: OC and SA	TTO: OC > SA (for 3L) DCE: OC \neq SA	$\Delta(S-O)$ and/or $\Delta(A-C)$
Mott et al. [28]	DCE	$\Delta(S-O)$	OC SC	Adults: OC Adolescents: SC	DCE: OC \neq SC	$\Delta(S-O)$ or sample com- position
Shah et al. [40]	TTO and DCE and VAS and LOD	$\Delta(S-O)$	OC SA	Adults: OC and SA	All methods: OC > SA	$\Delta(S-O)$ and/or $\Delta(A-C)$
Dewilde et al. [11]	TTO and VAS	$\Delta(S-O)$	OA OC	Adults: OA and OC	TTO: OC > OA VAS: OC > OA	$\Delta(A-C)$

[8, 12], EQ-5D-Y-3L health states are valued by adults using the perspective of a 10-year-old child¹ [35].

Several authors have demonstrated that the perspective used influences valuation of EQ-5D health states [11, 19, 21, 28, 40]. However, it remains unclear what drives these differences, which may (at least in part) be explained by the fact that the prescribed perspective for valuation of EQ-5D-Y-3L comprises not one but two primary changes from the perspective prescribed for health state valuation of adult EQ-5D instruments [34]. First, instead of adults valuing health states for their adult selves, they value health states for a child. This entails a change from an adult to a child perspective, which we denote as $\Delta(A-C)$. Second, instead of adults valuing health states for themselves, they value health states for someone else. This, in turn, entails a change from decision-making for the self to the other, which we denote as $\Delta(S-O)$. It has previously been hypothesized that both changes could separately influence decision-making in valuation tasks [26]. Disentangling the effect of these changes in perspective may require a factorial decomposition that combines $\Delta(A-C)$ and $\Delta(S-O)$, and hence provides an exploration of all possible combinations of the two primary changes. This decomposition yields four perspectives: self-adult (SA), other-adult (OA), self-child (SC), and other-child (OC).

Previous empirical research into valuation of EQ-5D-Y-3L has used only a subset of these four perspectives simultaneously for examining the impact of the perspective on EQ-5D-Y-3L valuation (see Table 1 for an overview of

the extant literature). Most of these studies appear to have selected methods and perspectives based on pragmatism and consistency with the methods used for adult EQ-5D valuation [35]. As a result, the studies presented in Table 1 predominantly used time trade-off (TTO) and discrete choice experiments (DCE), the preferred methods for valuation of EQ-5D [41]. All studies used the prescribed OC perspective for obtaining values for EQ-5D-Y-3L health states, some in combination with the SA perspective that is prescribed for valuation of adult EQ-5D health states [41]. Kreimeier et al. [21] find some evidence for higher valuation for the OC perspective compared to SA, using TTO and DCE. Shah et al. [40] reach similar conclusions for the worst possible health states described by EQ-5D-Y-3L, with the same methods. The latter also hold for visual analogue scales (VAS) and the location of dead (LOD) method [9]. The study designs applied by Kreimeier et al. [21] and Shah et al. [40], however, do not allow any conclusions on the relative contribution of $\Delta(S-O)$ and $\Delta(A-C)$ to (differences in) EQ-5D-Y-3L valuations. The only study that further decomposed the impact of $\Delta(S-O)$ and $\Delta(A-C)$ on valuations is that by Kind et al. [19], who used the SA, OA, and OC perspectives. Their results, however, are in contrast to those of others as Kind et al. [19] found evidence for lower valuation of child health states valued with VAS, which is an effect in the opposite direction to that observed by Kreimeier et al. [21] for TTO and by Shah et al. [40] for TTO and VAS. It could therefore be hypothesized that differences between TTO and VAS in the impact of the perspective may be explained by differences in willingness to trade-off life duration between perspectives, but this has not yet been explored.

In this study, we advance earlier work on the impact of the change in perspective on EQ-5D-Y-3L valuation, by

¹ Note that the exact instruction included in the protocol is: 'Considering your views about a 10-year-old child, what do you prefer?' (p. 658, [35]).

conducting a within-subjects experiment in which respondents completed TTO and VAS tasks from all four perspectives. Our study, thus, aims to provide additional insight into the way in which EQ-5D-Y-3L valuation depends on the perspective used. To facilitate a within-subjects design, we explored the impact of the four perspectives on EQ-5D-Y-3L valuation using TTO and VAS tasks, rather than the prescribed TTO and DCE tasks [35]. Furthermore, the use of VAS tasks enabled us to obtain further insight into the so far conflicting results for VAS found in the literature [19, 40].

The main objectives of this study were to explore (i) the extent to which the outcomes, precision, and quality of EQ-5D-Y valuation differ between the SA, OA, SC, and OC perspectives, (ii) what change(s) in perspective may drive any observed differences, and (iii) what influence the type of valuation task may have on any observed differences. Obtaining insight into the underlying mechanism driving differences in EQ-5D-Y-3L valuation may be of empirical interest and aid researchers in interpreting the results of valuation studies. Our results may also inform discussions on whether the change in prescribed perspective for valuation of EQ-5D-Y-3L health states—and the implications this change may have for estimation of QALY gains in health technology assessment—is considered desirable [26].

Methods

To meet the objective of our study, we designed a within-subjects experiment that was conducted using a sample of 205 bachelor students in Business Administration in September and October, 2020. The sample size was informed by a-priori power calculations, with the aim of at least being able to detect differences that classify as ‘small effects’ in a pairwise comparison between perspectives². Respondents were rewarded course credits for their participation, which lasted for 30 min. The sample consisted of 106 females (51.7%) and mean (SD) age was 19.48 (2.33). Respondents completed the experiment in individual cubicles, with an experienced researcher (SL or VRD) present to answer questions. The experiment was programmed in Shiny R (code available upon request).

Experimental procedure and design

The first part of the experiment was modelled in accordance with the EQ-5D-Y-3L valuation protocol [35]. As such, the experiment started with respondents completing

the EQ-5D-Y-3L instrument to familiarize themselves with its descriptive system, as well as reporting their age and sex. Next, the experiment commenced with a pre-recorded instruction video (see Online Supplements) in which the applied valuation tasks were explained by one of the researchers (SL)³. The respondents then valued EQ-5D-Y-3L health states across the severity range (see section “Health states” below) using the four perspectives: SA, OA, SC, and OC. We presented the health states and perspectives in random order to respondents. Respondents first valued the health states with VAS and subsequently with TTO (see section “Health state valuation” below). We did not randomize the order of VAS and TTO tasks to account for a possible learning effect that could affect the valuation of each health state, as TTO is often considered more difficult than VAS [5].

Health states

Health states were drawn from EQ-5D-Y-3L, which describes HRQOL using five dimensions: (i) mobility, (ii) looking after myself, (iii) usual activities, (iv) pain or discomfort and (v) feeling worried, sad, or unhappy. Problems on each dimension are described by three levels: having no problems (level 1), having some problems (level 2), or having a lot of problems (level 3). As such, the worst health state described by EQ-5D-Y-3L is, for example, coded as 33333. To examine the impact of the four perspectives across the severity range, we selected two blocks of health states from Kreimeier et al. [21] with the objective to include one mild, one moderate, and one severe health state. Furthermore, in both blocks we included 22222 and 33333 to enable testing for violations of logical consistency. Block 1 comprised the following health states: 11121, 22222, 32211, 33323, and 33333. Block 2 comprised: 11112, 11312, 22222, 13331, and 33333. Respondents were randomly assigned to one of these blocks.

Valuation methods

The written TTO and VAS task instruction for the four perspectives can be found in Table 2. VAS was operationalised by asking respondents to value health states using a slider that ranged from 0 to 100, with 0 representing ‘worst imaginable health’ and 100 representing ‘best imaginable health’. TTO was operationalised by a composite TTO procedure [17]. As such, each task introduced a life in impaired health, described as 10 years in a selected EQ-5D-Y-3L health state (Q_i), followed by immediate death. Valuation commenced

² Power-analysis based on a paired t-test with Cohen’s $d=0.2$, $\alpha=0.05$, $\beta=0.8$, i.e., using the parametrization proposed by Cohen (1988).

³ We opted for a pre-recorded video introduction to the valuation tasks to comply with the local COVID-19 regulations.

Table 2 TTO and VAS task instruction for each of four perspectives

	Self	Other
Child	Self-Child (SC) VAS: please rate the following health state for yourself as a 10-year-old child TTO: ‘Which life would be better for yourself as a 10-year-old child?’	Other-Child (OC) VAS: please rate the following health state for a 10-year-old child TTO: ‘Which life is better for a 10-year-old child?’
Adult	Self-Adult (SA) VAS: please rate the following health state for yourself TTO: ‘Which life is better for yourself?’	Other-Adult (OA) VAS: please rate the following health state for someone else (the same age as you) TTO: ‘Which life is better for someone else (the same age as you)?’

with a choice between 10 years in Q_i and immediate death. If respondents preferred living 10 years in Q_i over immediate death, valuation continued with a choice between 10 years in Q_i and 5 years in full health. By means of a bisection elicitation procedure with five choices in total, we obtained indifferences of the form: 10 years in $Q_i \sim X$ years in full health (with \sim denoting indifference). Under the assumption of a linear QALY model, such an indifference yielded the utility of health state Q_i as $X/10$ (Torrance [42]). If, however, respondents preferred immediate death over living 10 years in Q_i , an additional valuation task was required for valuation of this ‘worse-than-dead’ health state [10]. Valuation then continued with the choice between 5 years in full health and 10 years in full health followed by 10 years in Q_i . We then used the same bisection elicitation procedure to elicit respondent’s indifferences of the form: X years in full health \sim 10 years in full health followed by 10 years in Q_i . This yielded the utility of health state Q_i as $(X-10)/10$ [4, 10].

Statistical analyses and modelling strategy

We examined differences in valuation between the perspectives on three distinct levels, i.e., the outcomes, precision, and quality of the valuations. Note that we report the results without correction for multiple hypothesis testing considering that the objective of our study is of exploratory nature.

Valuation outcomes

The impact of the different perspectives on the outcomes of EQ-5D-Y-3L valuations was first explored by analysing within-subjects response patterns. These patterns illustrate the degree to which individual respondents valued the same state differently between the four perspectives, with a maximum of four unique valuations per health state (one for each

perspective). Furthermore, the direction of the impact of deciding for children and for others was explored by calculating individual-level differences between valuations from each perspective. Next, the overall impact of the different perspectives was investigated using linear mixed-effects regression models for all health states combined. Models were specified with subject random effects with the following fixed effects: (1) $\Delta(A-C)$, i.e., a dummy that distinguishes between the two adult (SA and OA) and child (SC and OC) perspectives, (2) $\Delta(S-O)$, i.e., a dummy that distinguishes between the two self (SA and SC) and other (OA and OC) perspectives, and (3) $\Delta(A-C) \times \Delta(S-O)$, i.e., an interaction term between these dummies that distinguishes the effect of deciding for another child (i.e., OC). Note that for these analyses we compiled the observations per valuation method and controlled for health state severity in one of two ways, resulting in four models. Models 1 and 2 report the results when controlling for health state severity by health state dummies (reference health state 11112), whereas models 3 and 4 report the results when controlling for health state severity by their level-sum-score (LSS), which is the sum of all problem level values for the five dimensions. That is, the reference health state 11112 has an LSS of 6 and 33333 has an LSS of 15. For ease of interpretation, we ran the regressions with LSS^r , i.e., $LSS^r = LSS - 6$. The latter models also allow testing of interactions between health state severity and $\Delta(A-C)$ or $\Delta(S-O)$.

Valuation precision

The impact of the different perspectives on the precision of valuations was explored by comparing variances between perspectives using a Bayesian modelling approach, operationalized as a JAGS model run in R. We opted for a Bayesian estimation process with non-informative priors, as this offers a more flexible and intuitive comparison of variances

that can take into account all observations of our within-subjects experiment simultaneously. Below, we present the general approach that was based on Golicki et al. [15], while the formal specification is available as Online Supplement. As Golicki et al. [15], we used a random-parameters model, which assumes that average (dis)utility associated with health states can differ between respondents. To estimate differences in standard deviation (SD) between perspectives we made the additional assumption that variances can differ between health states, as they are generally larger for more severe states [21, 49]. The impact of the perspectives on variance is modelled by two variance scaling factors (VSF) that capture $\Delta(A-C)$ and $\Delta(S-O)$. The VSF identifies differences in variance between adult and child perspectives as VSF-AC, with the VSF for $\Delta(S-O)$ denoted as VSF-SO. We used a Markov Chain Monte Carlo simulation with 500, 4000, and 20,000 adaptive, burn-in, and actual iterations, respectively, without thinning, and two chains. Besides reporting the median point estimate for the posterior distributions of VSF-AC and VSF-SO for VAS and TTO, we report the 2.5th and 97.5th percentiles to construct 95% credible intervals (CrI). We interpret a CrI that does not contain 1 as evidence for a difference in variance between perspectives. If the CrI for a scaling factor falls completely below (above) 1, this indicates that variances are smaller (larger).

Valuation quality

The impact of the different perspectives on the quality of valuations was explored following the approach used by Alava et al. [1]. As such, we categorised the following responses as problematic: (i) violations of dominance (e.g., assigning state 11121 a higher value than 22222), (ii) overall non-discrimination (i.e., assigning the same value to all health states within a perspective), and (iii) non-attendance (i.e., exiting the valuation task at the earliest point possible, which was 50 for VAS and $-0.5/0.5$ for TTO). Furthermore, we compared how the occurrence of three additional problems related to the VAS and TTO methods differed between perspectives. Specifically, we compared end-point usage for VAS (i.e., the use of 0 and 100 scores) as earlier work has identified that VAS may suffer from end-point aversion [43]. Finally, we compared the occurrence of non-trading responses (i.e., utilities of 1.0) and all-in trading (utilities of -1.0) for TTO, which may yield TTO data of problematic quality and/or lead to exclusion in other valuation studies due to outlying preferences [4, 13].

Results

Information about the means (and variance of) and distribution of VAS scores and TTO utilities per state and perspective can be found in the Online Supplements.

Valuation outcomes

Within subjects-effects

The observed within-subjects valuations differed substantially between perspectives. For example, we found that for VAS, the mean (SD) number of unique scores per respondent was 2.99 (0.94) out of four possible unique valuations per health state. The direction of the change varied between respondents. In some cases, these changes resulted in health states being considered better-than-dead when valued from one perspective but worse-than-dead from another. This, for example, occurred in 19% of respondents for state 33333. We further explored the impact of perspectives by determining the within-subjects effect of deciding for children and deciding for others. The following strategy was used for deciding for children: we calculated two difference scores per respondent per health state, i.e., (SA-SC) and (OA-OC). These difference scores may reflect the effect of deciding for children instead of adults within both self or both other perspectives respectively. A similar strategy may be used to investigate the effect of $\Delta(S-O)$ within subjects, i.e., by calculating (SA-OA) and (SC-OC). Figure 1 shows these difference scores, and indicates large heterogeneity between respondents completing VAS and TTO tasks for children rather than adults. That is, many respondents valued health states higher for children than for adults, while the opposite was also observed (see Online Supplements for a classification per state). Overall, the difference scores were moderately correlated for both methods [VAS: $r(1018)=0.51$, $p<0.001$; TTO: $r(1018)=0.38$, $p<0.001$], indicating that the observed heterogeneity was systematic (correlations per health state can be found in Online Supplements). This systematicity reflects that although the direction of the impact of deciding for children differed between respondents, it was often consistent within-subjects. In other words, if a respondent valued a health state higher for a child than an adult in both self-perspectives (SA-SC < 0), the same state was most likely also valued higher for a child in both other-perspectives (OA-OC < 0). If the within-subjects effect of deciding for children was in the opposite direction, this consistency held (i.e., SA-SC > 0 and OA-OC > 0). With regard to deciding for others, we observed a weak correlation between differences scores for TTO ($r(1018)=0.13$, $p<0.001$), but not for VAS ($r(1018)=0.00$).

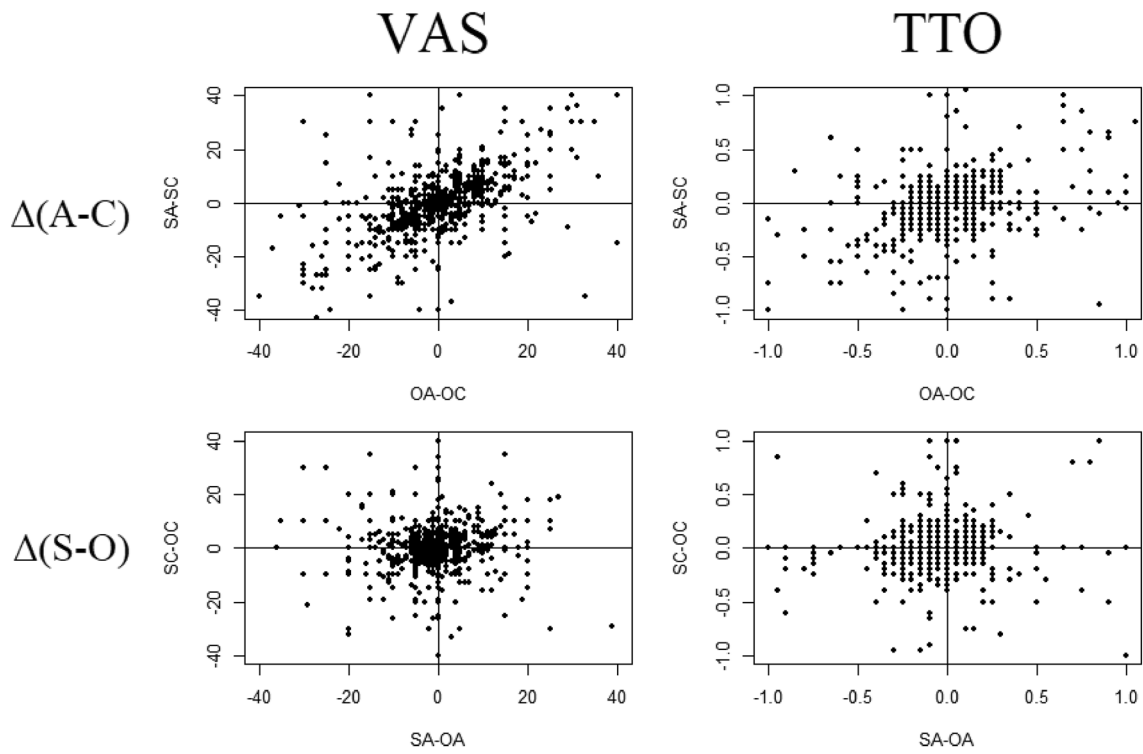


Fig. 1 Within-subjects differences between perspectives for VAS (left) and TTO (right). Upper panels show differences between child and adult perspectives ($\Delta(A-C)$) and lower panels show differences between self and other perspectives $\Delta(S-O)$

Fig. 2 Mean VAS scores per EQ-5D-Y health state and perspective

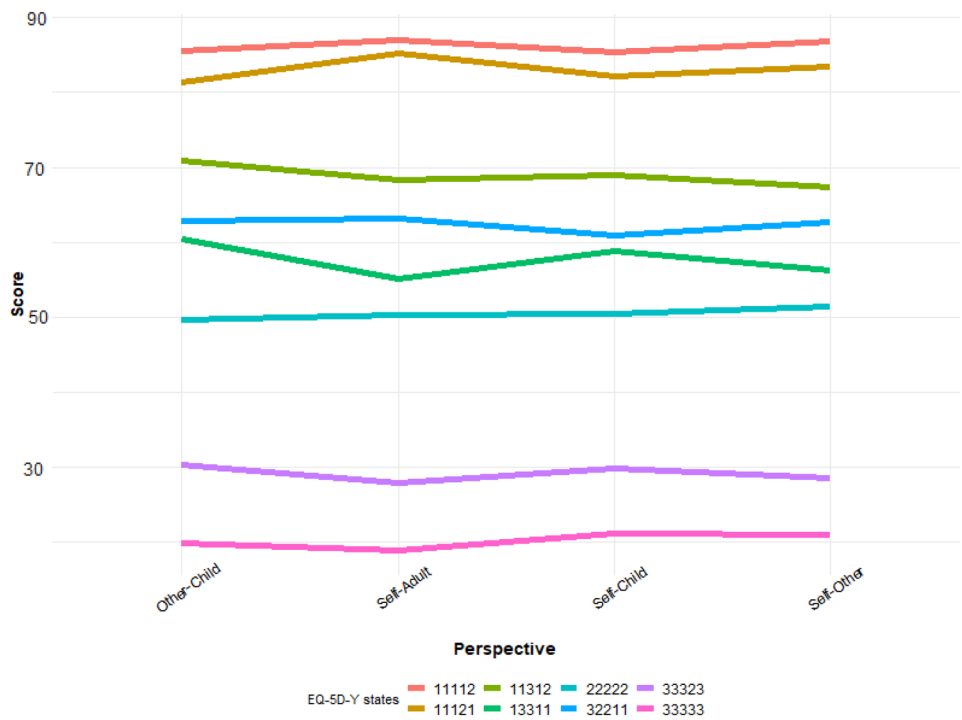
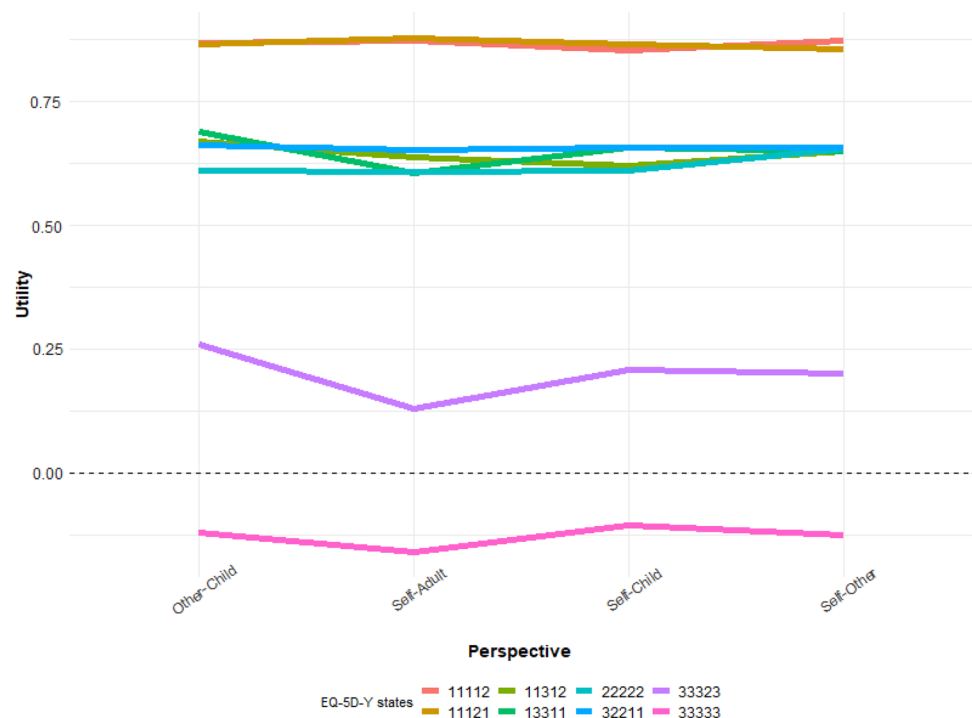


Fig. 3 Mean TTO utilities per EQ-5D-Y health state and perspective



Overall effects

Figures 2 and 3 show the impact of perspective on valuation outcomes for VAS and TTO, respectively. Although our within-subjects analyses suggest that considerable heterogeneity exists, overall, the differences between perspectives are generally small to non-existent between health states. In those cases where we did observe (larger) differences, their directions depended on the health state. This is substantiated by a set of regression analyses per health state, reported in the Online Supplements. The results of the linear mixed effects regressions, reported in Table 3, show that, controlling for health states, deciding for others yielded higher TTO utilities (compared to deciding for oneself). However, this was not observed when controlling for LSS (Models 3 and 4). Instead, a positive interaction between LSS and $\Delta(A-C)$ was observed, which suggests that although severe health states receive lower valuations, this effect may be less pronounced for children than for adults.

Valuation precision

Table 4 shows the results of the Bayesian approach used to estimate differences in variance. For VAS, the posterior distribution for VSF-AC suggested that variances were larger for adult perspectives than for child perspectives. For TTO, variances were similar between child and adult perspectives, but the variance of TTO utilities was smaller when respondents valued health states for other adults as compared to for themselves.

Valuation quality

Table 5 provides an overview of the quality indicators per perspective. A higher count indicates lower data quality, except for the indicator 'end-point usage', where a higher count indicates higher quality data. The proportion of problematic responses ranged from 0 to 15% of responses between categories. Our results suggest that violations of dominance were not independently distributed, yet occurred more frequently for child perspectives (in particular, for OC). Respondents were also more likely to use the end-point of the VAS for children. Finally, non-trading was more likely for respondents deciding for themselves than for children.

Discussion

In this study, we explored the impact of the perspective on valuation of EQ-5D-Y-3L health states in a sample of students, by comparing the decomposed influence of two primary changes: (i) deciding for an adult or a child, and (ii) deciding for oneself or another person. Overall, the results of our study show that the four decomposed perspective impact the outcomes, precision, and quality of EQ-5D-Y-3L valuation differently. The overall impact appears to be small. Nonetheless, our work provides detailed insight into what drives differences between similar EQ-5D-Y-3L health states for children and adults that can be found in the literature. Our results suggest that the effect of perspective is highly heterogeneous between health states and respondents. That

Table 3 Fixed effects estimates (standard errors in brackets) for mixed effects regression analyses for VAS scores and TTO utilities

Model	Severity approach 1: health state dummies		Severity approach 2: LSS	
	1	2	3	4
Method	VAS	TTO	VAS	TTO
Intercept	85.74 (1.34)***	0.84 (0.02)***	82.86 (1.28)***	0.94 (0.02)***
Fixed effects				
$\Delta(A-C): C$	0.47 (0.75)	0.02 (0.01)	- 1.24 (1.23)	- 0.02 (0.02)
$\Delta(S-O): O$	0.46 (0.75)	0.03 (0.01)*	- 0.72 (1.23)	0.01 (0.02)
$\Delta(S-O) \times \Delta(A-C): OC$	- 0.32 (1.05)	- 0.02 (0.02)	1.56 (1.74)	0.01 (0.03)
LSS ^r			- 7.25 (0.17)***	- 0.11 (0.003)
LSS ^r $\times \Delta(A-C): C$			0.42 (0.24)+	0.01 (0.004)*
LSS ^r $\times \Delta(S-O): O$			0.29 (0.24)	0.00 (0.00)
LSS ^r $\times \Delta(S-O) \times \Delta(A-C): OC$			- 0.46 (0.33)	- 0.01 (0.01)
HS: 11121	- 2.95 (1.42)*	0.00 (0.03)		
HS: 11312	- 17.33 (1.10)***	- 0.22 (0.02)***		
HS: 13311	- 28.54 (1.10)***	- 0.22 (0.02)***		
HS: 22222	- 35.70 (1.03)***	- 0.25 (0.03)***		
HS: 32211	- 23.63 (1.42)***	- 0.21 (0.03)***		
HS: 33323	- 56.88 (1.42)***	- 0.67 (0.03)***		
HS: 33333	- 65.92 (1.03)***	- 0.99 (0.02)***		

LSS^r = rescaled level-sum-score; ***, **, * and + indicate (marginal) significance at $p < 0.001$, $p < 0.01$, $p < 0.05$ and $p < 0.10$, respectively

Table 4 Medians and 95% CrI for scaling factors that identify between-subjects differences in variance

Method	Parameter	Median	95% CrI
VAS	VSF-AC	1.053	[1.008, 1.101]
	VSF-SO	0.970	[0.928, 1.013]
TTO	VSF-AC	1.043	[0.995, 1.092]
	VSF-SO	0.902	[0.861, 0.942]

is, deciding for children and deciding for others yielded higher valuations for some health states and respondents, whereas the opposite was also observed. As such, the results of this study may provide an explanation for some of the (sometimes conflicting) findings in other work comparing the use of different perspectives in EQ-5D-Y-3L with VAS and TTO.

Our findings for VAS valuation outcomes suggest that VAS scores may differ systematically when elicited from an adult or child perspective, but that differences could potentially be small and occur in either direction. These results both align and conflict with earlier work. That is, Kind et al. [19] found a near-uniform pattern of lower valuations for children. Other work by Dewilde et al. [11], found that when adults decide for another adult or for a 10-year-old child, VAS scores are generally lower for adults. Similarly, Shah et al. [40] found that state 33333 is valued higher

for children. Our work supports this ambiguity, as we find effects in both directions. For TTO, our results on the effect of deciding for children or adults are similarly mixed. Although some authors have found evidence of higher TTO utilities for children than for adults [11, 40], existing work is also not conclusive about the effect of deciding for children rather than for adults. For example, Kreimeier et al. [21] found that differences between SA and the OC perspective occurred in both directions depending on the health state valued, similar to those observed in the current study.

The heterogeneity in valuations for children, observed in both our study and the extant literature, could have different explanations. Dewilde et al. [11], for example, identified conflicting beliefs about the impact of health impairment on adults and children in a think-out-loud study. Some respondents believed adults were better able to cope with any impairment than children, while others believed the opposite to be true. Other conflicting arguments provided were the importance of adult health to be able to take care of children, as well as the importance of childhood as a foundational period for further development. This suggests that the influence may depend on who is valuing the health state. Our results also show a trend suggesting a relationship between the severity of the health state considered and differences between health states valued for adults and children—suggesting that differences depend on and may increase with severity. Hence, differences between studies could be related

Table 5 Quality indicators per perspective (with the maximum possible violations *per perspective* in brackets)

Quality indicator	Self-adult	Other-adult	Self-child	Other-child
Dominance violation (max. $n = 1406$)				
VAS*	101	102	129	214
TTO*	79	68	92	208
Overall non-discrimination (max $n = 205$)				
VAS	0	0	1	0
TTO	0	0	2	0
Non-attendance (max $n = 1025$)				
VAS	51	56	46	32
TTO	93	106	95	85
End-point usage (max $n = 1025$)				
VAS*	65	45	89	81
Non-trading responses (max $n = 1025$)				
TTO*	62	46	79	63
All-in trading responses (max $n = 1025$)				
TTO	37	27	33	27

*Indicates that the distributions was not independent between perspectives, Chi-squared tests, $p < 0.05$

to the states selected. However, this should be substantiated in future work.

Few studies have compared the effect of deciding for others and deciding for oneself. Our study showed that for VAS this could have effects in either direction. Hence, our results are similar to those of Kind et al. [19], who found evidence for differences in VAS scores between adults deciding for themselves (SA) or others (OA) for a small number of health states, but the direction of these effects differed between countries and health states. For TTO, our results suggest that deciding for others may have a small but upward effect on utilities overall, which to our knowledge is a novel result. This implies that respondents were less willing to give up life years for others. Future work could explore why this occurs. For example, it may be the case that individuals discount others' life years less (and are thus less willing to give up others' years in the future). There is some evidence suggesting such differences in time preferences for self and others [2, 32, 38, 52], but these findings typically involve monetary outcomes and are inconclusive about the direction of these differences. Alternatively, some respondents may feel hesitant to give up others' life years as they feel it is not their choice to make. Earlier work has found effects in a similar direction related to religion and beliefs in life after death [16, 48], which could reflect a similar position held by individuals trading off their own health.

Our study also allowed exploring valuation precision and quality between perspectives. Although such analyses may be relevant in the context of EQ-5D-Y-3L valuation, they have not extensively been reported in earlier work comparing perspectives for valuation of EQ-5D-Y-3L (an exception is: [28]). Our findings suggested that more violations of dominance occurred when health states were valued for children

rather than adults, which is analogous to the findings reported for data quality in [28]. Our work is, to our knowledge, the first to report lower variance observed for EQ-5D health states valued for someone else than for oneself in TTO. This suggests that respondent's decisions about health states are more similar when deciding for others than when they deciding for themselves. A possible explanation can be found in construal level theory [44]. This theory states that psychological distance affects whether individual's thought processes are concrete or abstract. Hence, for themselves (low psychological distance), individuals are more likely to take into account their own concrete situation but for others (high psychological distance) people focus less on details. However, we also find higher variance for VAS valuation from adult perspectives, which seems to be in conflict with this explanation. Hence, future work may further explore the causes and implications of variance in valuations.

Overall, the observed differences between perspectives may be considered small in comparison to those observed by others [11, 40]. Seeing as in our study the differences observed were smaller by an order of magnitude (e.g., the regression coefficient for $\Delta(S-O)$ was 0.03 for TTO), this raises the question whether these differences are meaningful. We do for three reasons. First, it may help to emphasize that a difference of, for example, 0.03 is the result of individuals trading off 3% more of their remaining lifetime in TTO. Even though these are hypothetical questions, we would caution against classifying such a sacrifice as trivial. The fact that many individuals are loss averse and would give up these life years reluctantly underlines non-triviality [6, 24]. Second, the differences observed in our study are in line with many of the estimated minimally important differences for EQ-5D [27, 50], suggesting that differences of this

magnitude may at least be of clinical relevance. Third, the observed differences are of a magnitude that could be seen as practically relevant for decision makers, as the median incremental QALY gain in published cost-effectiveness studies was estimated to be of similar size Wisløff et al. [51].

The discrepancies (in direction and magnitude) between our results and those of others may further be related to the following limitations of our study. First, in most of the extant studies, EQ-5D-Y-3L valuation was completed in one-to-one personal interviews facilitated by a trained interviewer [11, 21, 40]. Our study, instead, asked respondents to work through these tasks by themselves after receiving a video instruction. Some evidence exists that individuals completing these tasks without supervision (i.e., online) may yield data of lower quality [31]. Unfortunately, our experimental set-up, as well as restrictions relating to the COVID-19 pandemic, precluded the use of one-to-one personal interviews in the lab or the use of digital interviewer-assisted interviews [23]. Note that we aimed to compensate for this by having an experienced researcher present at all times. However, our analyses of data quality suggest that the data obtained from this study is of reasonable quality. For example, Ramos-Goñi et al. [36] report that even with personal interviews and extensive quality control, 19% (25%) of Dutch (Spanish) respondents had at least one inconsistent response in EQ-5D-5L valuation. Alava et al. [1] investigated the quality of the data reported in Devlin et al. [8], which who did not employ quality control, and found over 90% to have at least one inconsistent violation. In our sample, that proportion for SA perspectives was 27%. Second, this study used a sample of students, whereas valuation studies use samples representative of the general adult population [11, 21, 28, 40]. Hence, the small differences observed between child and adult perspectives may be explained by the fact that our sample consisted of respondents who are still transitioning into adulthood. While students may remember their recent childhood better than older adults, making it easier for them to value health states for a child perspective, they will likely not (yet) have any children of their own, which may be relevant for EQ-5D-Y-3L valuation. It is also well known that some of the demographics in which students differ from the general public may affect health state valuation, such as age and education level [8, 25, 47]. Furthermore, even though the student sample included a students of many nationalities, there is evidence of Dutch EQ-5D valuation having different characteristics [33]. Hence, it is recommended that future work considers to replicate our approach in general public samples. Third, the bisection choice procedure implemented in this study differs from the elicitation procedure used or recommended by others [21, 37]. Although this change facilitated self-completion, it is well known that the elicitation procedure can influence TTO results [3]. Fourth, although we aimed for both health

state blocks to be assigned randomly, the final distribution was imbalanced suggesting randomisation error could have taken place. Finally, in order to increase power, we opted for a within-subjects approach, in which the order of each perspective was randomized. Although such an approach will help identify within-subjects effects and differences, it may also be sensitive to order effects and perhaps anchoring [45]. That is, individuals may have anchored on their initial answer for a single perspective and adjusted their answers for subsequent perspectives insufficiently. Although randomizing the order ensures this does not systematically bias our results on a sample level, anchoring could explain why the differences in outcomes between perspectives appear smaller than in some of the findings published in earlier work. Future research could test this hypothesis using the four perspectives we used in a between-subjects experiment.

To conclude, our study showed that the use of different perspectives will likely yield (at least) small, but systematic differences in the outcomes, precision, and quality of valuation of EQ-5D-Y-3L health states. Our exploration of the causes of these effects suggested that TTO utilities are affected in upward direction when one is asked to decide for others (rather than for oneself). Deciding for children (rather than adults) can affect EQ-5D-Y-3L valuation with both VAS and TTO, but the direction of this effect was ambiguous and differed between respondents and health states in our study. If these results generalise to the samples used in EQ-5D-Y-3L valuation, the small observed differences may have consequences for estimation of QALY gains [26]. Nonetheless, the large heterogeneity between respondents and states we found suggests that the search for the empirical and normative implications of perspective used in EQ-5D-Y-3L is far from over.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10198-021-01377-y>.

Acknowledgements This study was made possible through funding provided by the EuroQol Research Foundation (project number: 20190890). The views expressed by the authors do not necessarily reflect the views of the EuroQol Group. Michal Jakubczyk and Milad Karimi are members of the EuroQol group.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Alava, M.H., Pudney, S., Wailoo, A.: The Eq-5d-5l value set for England: findings of a quality assurance program. *Value Health* (2020). <https://www.sciencedirect.com/science/article/pii/S1098301520300838>
2. Albrecht, K., Volz, K.G., Sutter, M., Laibson, D.I., von Cramon, D.Y.: What is for me is not for you: brain correlates of intertemporal choice for self and other. *Soc. Cogn. Affect. Neurosci.* **6**, 218–225 (2011)
3. Attema, A.E., Edelaar-Peeters, Y., Versteegh, M.M., Stolk, E.A.: Time trade-off: one methodology, different methods. *Eur. J. Health Econ.* **14**, 53–64 (2013)
4. Attema, A.E., Versteegh, M.M., Oppe, M., Brouwer, W.B., Stolk, E.A.: Lead time TTO: leading to better health state valuations? *Health Econ.* **22**, 376–392 (2013)
5. Badia, X., Monserrat, S., Roset, M., Herdman, M.: Feasibility, validity and test-retest reliability of scaling methods for health states: the visual analogue scale and the time trade-off. *Qual. Life Res.* **8**, 303–310 (1999)
6. Bleichrodt, H.: A new explanation for the difference between time trade-off utilities and standard gamble utilities. *Health Econ.* **11**, 447–456 (2002). <https://personal.eur.nl/bleichrodt/he2002.pdf>
7. Chen, G., Ratcliffe, J.: A review of the development and application of generic multi-attribute utility instruments for paediatric populations. *Pharmacoeconomics* **33**, 1013–1028 (2015)
8. Devlin, N.J., Shah, K.K., Feng, Y., Mulhern, B., van Hout, B.: Valuing health-related quality of life: an eq-5 d-5 l value set for England. *Health Econ.* **27**, 7–22 (2018)
9. Devlin, N.J., Shah, K.K., Mulhern, B.J., Pantiri, K., van Hout, B.: A new method for valuing health: directly eliciting personal utility functions. *Eur. J. Health Econ.* **20**, 257–270 (2019)
10. Devlin, N.J., Tsuchiya, A., Buckingham, K., Tilling, C.: A uniform time trade off method for states better and worse than dead: feasibility study of the 'lead time' approach. *Health Econ.* **20**, 348–361 (2011)
11. Dewilde, S., Janssen, M.F., Lloyd, A.J., K.K., S.: Exploration of the reasons why health state valuation differs for children compared to adults: a mixed methods approach. working paper (2020)
12. Dolan, P., Gudex, C., Kind, P., Williams, A.: The time trade-off method: results from a general population study. *Health Econ.* **5**, 141–154 (1996)
13. Engel, L., Bansback, N., Bryan, S., Doyle-Waters, M.M., Whitehurst, D.G.: Exclusion criteria in national health state valuation studies: a systematic review. *Med. Decis. Making* **36**, 798–810 (2016)
14. EuroQol Research Foundation. EQ-5D-Y User Guide, 2020. <https://euroqol.org/publications/user-guides>
15. Golicki, D., Jakubczyk, M., Graczyk, K., Niewada, M.: Valuation of Eq-5d-5l health States in Poland: the first Eq-Vt-based study in central and Eastern Europe. *Pharmacoeconomics* **37**, 1165–1176 (2019)
16. Jakubczyk, M., Golicki, D., Niewada, M.: The impact of a belief in life after death on health-state preferences: true difference or artifact? *Qual. Life Res.* **25**, 2997–3008 (2016)
17. Janssen, B.M., Oppe, M., Versteegh, M.M., Stolk, E.A.: Introducing the composite time trade-off: a test of feasibility and face validity. *Eur. J. Health Econ.* **14**, 5–13 (2013)
18. Kennedy-Martin, M., Slaap, B., Herdman, M., Van Reenen, M., Kennedy-Martin, T., Greiner, W., Busschbach, J., Boye, K.S.: Which multi-attribute utility instruments are recommended for use in cost-utility analysis? A review of national health technology assessment (HTA) guidelines. *Eur. J. Health Econ.* **21**, 1245–1257 (2020). <https://doi.org/10.1007/s10198-020-01195-8>
19. Kind, P., Klose, K., Gusi, N., Olivares, P.R., Greiner, W.: Can adult weights be used to value child health states? testing the influence of perspective in valuing Eq-5d-Y. *Qual. Life Res.* **24**, 2519–2539 (2015)
20. Kreimeier, S., Greiner, W.: Eq-5d-Y as a health-related quality of life instrument for children and adolescents: the instrument's characteristics, development, current use, and challenges of developing its value set. *Value in Health* **22**, 31–37 (2019)
21. Kreimeier, S., Oppe, M., Ramos-Goñi, J.M., Cole, A., Devlin, N., Herdman, M., Mulhern, B., Shah, K.K., Stolk, E., Rivero-Arias, O.: 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 Health* **21**, 1291–1298 (2018)
22. Kwon, J., Kim, S.W., Ungar, W.J., Tsiplova, K., Madan, J., Petrou, S.: Patterns, trends and methodological associations in the measurement and valuation of childhood health utilities. *Qual. Life Res.* **28**, 1705–1724 (2019)
23. Lipman, S.A.: Time for tele-Tto? Lessons learned from digital interviewer-assisted time trade-off data collection. *Patient* **14**, 459–469 (2020). <https://repub.eur.nl/pub/132631>
24. Lipman, S.A., Brouwer, W.B.F., Attema, A.E.: Qalys without bias? Non-parametric correction of time trade-off and standard gamble weights based on prospect theory. *Health Econ.* **28**, 843–854 (2019)
25. Lipman, S.A., Brouwer, W.B.F., Attema, A.E.: Living up to expectations: experimental tests of subjective life expectancy as reference point in time trade-off and standard gamble. *J. Health Econ.* **71**, 102318 (2020)
26. Lipman, S.A., Reckers-Droog, V.T., Kreimeier, S.: Think of the children: a discussion of the rationale for and implications of the perspective used for Eq-5d-Y health state valuation. *Value Health* **24**, 976–982 (2020)
27. McClure, N.S., Sayah FA, Xie F, Luo, N., Johnson J.A.: Instrument-defined estimates of the minimally important difference for Eq-5d-5l index scores. *Value Health* **20**, 644–650 (2017). <https://doi.org/10.1016/j.jval.2016.11.015>
28. Mott, D., Shah, K., Ramos-Goñi, J., Devlin, N., Rivero-Arias, O.: Valuing Eq-5d-Y health states using a discrete choice experiment: do adult and adolescent preferences differ? (2019). <https://doi.org/10.1177/0272989X21999607>
29. Mulhern, B., Bansback, N., Brazier, J., Buckingham, K., Cairns, J., Devlin, N., Dolan, P., Hole, A. R., Kavetsos, G. and Longworth, L.: Preparatory study for the revaluation of the Eq-5d tariff: methodology report. *Health Technol. Assess.*, **18**. (2014)
30. NICE.: Guide to the processes of technology appraisal. In: Excellence, N. I. F. H. A. C. (Ed). London. (2018)
31. Norman, R., King, M.T., Clarke, D., Viney, R., Cronin, P., Street, D.: Does mode of administration matter? Comparison of online and face-to-face administration of a time trade-off task. *Qual. Life Res.* **19**, 499–508 (2010)
32. O'Connell, G., Christakou, A., Haffey, A.T., Chakrabarti, B.: The role of empathy in choosing rewards from another's perspective. *Front. Hum. Neurosci.* **7**, 174 (2013)
33. Olsen, J.A., Lamu, A.N., Cairns, J.: In search of a common currency: a comparison of seven Eq-5d-5l value sets. *Health Econ.* **27**, 39–49 (2018)
34. Oppe, M., Devlin, N.J., van Hout, B., Krabbe, P.F., de Charro, F.: A program of methodological research to arrive at the new international Eq-5d-5l valuation protocol. *Value Health* **17**, 445–453 (2014)
35. Ramos-Goni, J., Oppe, M., Stolk, E., Rivero-Arias, O., Devlin, N., Shah, K., Kreimeier, S.: The international valuation protocol for the Eq-5d-Y-3l. *Pharmacoeconomics* **38**, 653–663 (2020)

36. Ramos-Goñi, J.M., Oppe, M., Slaap, B., Busschbach, J.J., Stolk, E.: Quality control process for Eq-5d-5l valuation studies. *Value Health* **20**, 466–473 (2017)
37. Ramos-Goñi, J. M., Oppe, M., Stolk, E., Shah, K., Kreimeier, S., Rivero-Arias, O. and Devlin, N.: International valuation protocol for the Eq-5d-Y-3l. *Pharmacoeconomics*. 1–11. (2020)
38. Rau, H.A.: Time preferences in decisions for others. *cege Discussion Papers*. (2020). <https://www.econstor.eu/bitstream/10419/219024/1/1700118161.pdf>
39. Rowen, D., Rivero-Arias, O., Devlin, N., Ratcliffe, J.: Review of valuation methods of preference-based measures of health for economic evaluation in child and adolescent populations: where are we now and where are we going? *Pharmacoeconomics* **38**, 325–340 (2020)
40. Shah, K. K., Ramos-Goñi, J. M., Kreimeier, S., Devlin, N. J. An Exploration of Methods for Obtaining 0= Dead Anchors for Latent Scale Eq-5d-Y Values. *Eur. J. Health Econ.* 1–13. (2020)
41. Stolk, E., Ludwig, K., Rand, K., van Hout, B., Ramos-Goñi, J.M.: Overview, update, and lessons learned from the international eq-5d-5l valuation work: version 2 of the Eq-5d-5l valuation protocol. *Value Health* **22**, 23–30 (2019)
42. Torrance, G.W.: Utility approach to measuring health-related quality of life. *J. Chronic Dis.* **40**(6), 593–600 (1987). [https://doi.org/10.1016/0021-9681\(87\)90019-1](https://doi.org/10.1016/0021-9681(87)90019-1)
43. Torrance, G.W., Feeny, D., Furlong, W.: Visual analog scales: do they have a role in the measurement of preferences for health states? *Med. Decis. Making* **21**, 329–334 (2001)
44. Trope, Y., Liberman, N.: Construal-level theory of psychological distance. *Psychol. Rev.* **117**, 440 (2010)
45. Tversky, A., Kahneman, D.: Judgment under uncertainty: heuristics and biases. *Science* **185**, 1124–1131 (1974)
46. UNGAR, W.J. : Challenges in health state valuation in paediatric economic evaluation. *Pharmacoeconomics* **29**, 641–652 (2011)
47. Van Nooten, F.: The person behind the Tto: exploring what influences Tto responses. (2019). <https://repub.eur.nl/pub/115653/>
48. van Nooten, F., van Exel, N., Eriksson, D., Brouwer, W.: “Back to the future”: influence of beliefs regarding the future on tto answers. *Health Qual. Life Outcomes* **14**, 4 (2016)
49. Versteegh, M.M., Vermeulen, K.M., Evers, S.M., de Wit, G.A., Prenger, R., Stolk, E.A.: Dutch tariff for the five-level version of Eq-5d. *Value Health* **19**, 343–352 (2016)
50. Walters, S.J., Brazier, J.E.: Comparison of the minimally important difference for two health state utility measures: Eq-5d and Sf-6d. *Qual. Life Res.* **14**, 1523–1532 (2005). <https://doi.org/10.1007/s11136-004-7713-0>
51. Wisløff, T., Hagen, G., Hamidi, V., Movik, E., Klemp, M., Olsen, J.A.: Estimating Qaly gains in applied studies: a review of cost-utility analyses published in 2010. *Pharmacoeconomics* **32**, 367–375 (2014)
52. Ziegler, F.V., Tunney, R.J.: Decisions for others become less impulsive the further away they are on the family tree. *PLoS One* **7**, e49479 (2012). <https://doi.org/10.1371/journal.pone.0049479>
53. ZINL.: Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg. Diemen: Zorginstituut Nederland. (2015)

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