Centre for Economic Performance

Occasional Paper

No.69
April 2025

The C-WELLBY: Towards a Universal Measure of Children's Wellbeing for Policy Analysis

Isaac Parkes

LSE THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

UKRI Economic and Social Research Council

Abstract
There is a pressing need for a universal measure of children's wellbeing, parallel to the WELLBY, for use in cost-benefit analysis. Currently, there is no consensus on how to value the wellbeing of children, raising concerns that their welfare is being undervalued in policy decisions. In this report, we discuss the issues inherent in measuring children's wellbeing and investigate a policy-oriented solution: the C-WELLBY.

Children aged 10 and above generally demonstrate stable, valid responses to evaluative life satisfaction questions, as evidenced by analyses of both Understanding Society and the Active Lives of Children and Young People Survey. Consequently, we recommend the use of WELLBYs, valued at the usual £15,920 in 2024 prices (HMT, 2021), for cost-benefit analyses of policy affecting this age group. For children aged below 10, we recommend estimating a C-WELLBY using mapping functions, also valued at £15,920.

Keywords: wellbeing, children, mental health, life-satisfaction, measurement.
JEL: H43, H75, I140, I20, I30

Isaac Parkes, Centre for Economic Performance at LSE.

This paper provides background analysis to Value for Money: how to increase wellbeing and reduce misery (D. Frayman, C.Krekel, R. Layard, S. MacLennan and I. Parkes). CEP Special Report 44. (2024).

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means without the prior permission in writing of the publisher nor be issued to the public or circulated in any form other than that in which it is published.

Requests for permission to reproduce any article or part of the Occasional Paper should be sent to the editor at the above address.

I. Parkes, submitted April 2025.

The C-WELLBY: Towards a Universal Measure of Children's Wellbeing for Policy Analysis[1]

Parkes, I.

London School of Economics

Executive Summary

Introduction

There is a pressing need for a universal measure of children's wellbeing, parallel to the WELLBY, for use in cost-benefit analysis. Currently, there is no consensus on how to value the wellbeing of children, raising concerns that their welfare is being undervalued in policy decisions. In this report, we discuss the issues inherent in measuring children's wellbeing and investigate a policy-oriented solution: the C-WELLBY.

When to use C-WELLBYs

Children aged 10 and above generally demonstrate stable, valid responses to evaluative life satisfaction questions, as evidenced by analyses of both Understanding Society and the Active Lives of Children and Young People Survey. Consequently, we recommend the use of WELLBYs, valued at the usual £15,920 in 2024 prices (HMT, 2021), for cost-benefit analyses of policy affecting this age group. For children aged below 10, we recommend estimating a C-WELLBY, also valued at £15,920.

Estimating C-WELLBYs

We evaluate the use of mapping functions to estimate a child's life satisfaction (a C-WELLBY), based on two alternative measures of wellbeing: the Strengths and Difficulties Questionnaire (SDQ) and happiness (affective wellbeing) scores. Evidently, where children below 10 cannot reliably answer life satisfaction questions, it is impossible to estimate mapping functions for this age group. Instead, we estimate this relationship for those aged 11-15 and discuss the feasibility of extrapolating our results below age 10.

[1] I am grateful to Sara MacLennan, Richard Layard, Christian Krekel and Ekaterina Oparina at the Centre for Economic Performance (LSE) for their continued help with this paper. I am deeply thankful to Allan Little for his pioneering role in the conception of the C-WELLBY. I am finally thankful to Clara Mukuria, Tamsin Ford, Eleanor Rees, Jessica Deighton, Jon Franklin, Qiqi Cheng, and Delphine Chollet for helpful discussions on this paper.

Analyses of both Understanding Society and the Active Lives of Children and Young People Survey data show that mapping can successfully predict life satisfaction from the SDQ and Happiness scores for those age 11-15. There is minimal variation in the estimated mapping functions by age, providing suggestive evidence that extrapolating to younger children is feasible, although further research is needed to ensure external validity and refine the accuracy of this approach.

Using C-WELLBYs

At this stage, we propose the use of a linear mapping function of the SDQ Total Difficulties Score to estimate a C-WELLBY for children below 10, as it provides a clear, simple translation. This estimated coefficient is -0.146, so:

- If an intervention decreases average SDQ score for one year by 2 points, then the estimated corresponding increase in average life satisfaction would be -2 * -0.146 = **0.292 C-WELLBYs**.
- Or a monetised wellbeing benefit of £15,920 * 0.292 = **£4,649**.

We also propose the mapping of momentary happiness scores in the same manner for this age group, using an estimated coefficient of 0.546.

For those aged 10-15, we recommend that the question "How do you feel about life as a whole? Not at all happy - Completely happy" asked in the UK Household Longitudinal Study 'Understanding Society' youth panel, is considered equivalent to life satisfaction[2].

Discussion

A single universal measure for children's wellbeing remains elusive. For younger children (2-9), a proxy method of some sort is necessary. Using mapping functions – common in health economics – offers a promising route to a uniform measure within the WELLBY framework. Future research would benefit from the creation of lookup tables with more complex functional forms and the external validation of these mapping functions. In the interim, the use of a simple, linear mapping coefficient for the SDQ (-0.146) or for happiness (0.546) is recommended to ensure children's wellbeing is appropriately valued in policy.

[2] See section 2 for more details.

1. Introduction:

There is a pressing need for a universal measure of children's wellbeing. The WELLBY provides a simple, monetizable measure of adult wellbeing for cost-benefit analysis. However, WELLBYs are not appropriate for young children. Despite a plethora of validated measures of children's wellbeing, none have stood out as a clear counterpart to life satisfaction and WELLBYs. This is because children below a certain age cannot reliably respond to evaluative life satisfaction questions; currently the only monetizable subjective wellbeing criterion in the HM Treasury's Green Book (HMT, 2021). Hence, there are calls for the introduction of a corresponding wellbeing measure for children to the Green Book, allowing the inclusion of wellbeing effects for analysis of policy affecting school-age children. This is the C-WELLBY.

The concept originated with the Department for Education (DfE), who in September 2022 notionally described the "C-WELLBY" as a WELLBY-equivalent summary measure of wellbeing and/or emotional health during childhood, for use in economic appraisal. This concept was further developed through a discussion paper by Little and Parkes (2024), which established key foundational concepts for mapping child wellbeing measures to the WELLBY and proposed applying the Treasury's WELLBY value to children on equity grounds.

Building on these origins, this paper further evaluates potential solutions to what Little and Parkes termed the 'missing childhood problem'. Section 2 asserts the validity of evaluative wellbeing measures for children aged 10 and over. Section 3 outlines the conceptual problems with children's wellbeing measurement below 10 and discusses possible solutions. Section 4 provides an example of how a solution using mapping functions could be enacted. Section 5 concludes and provides recommendations for policy and further research.

2. Age 10 and Over

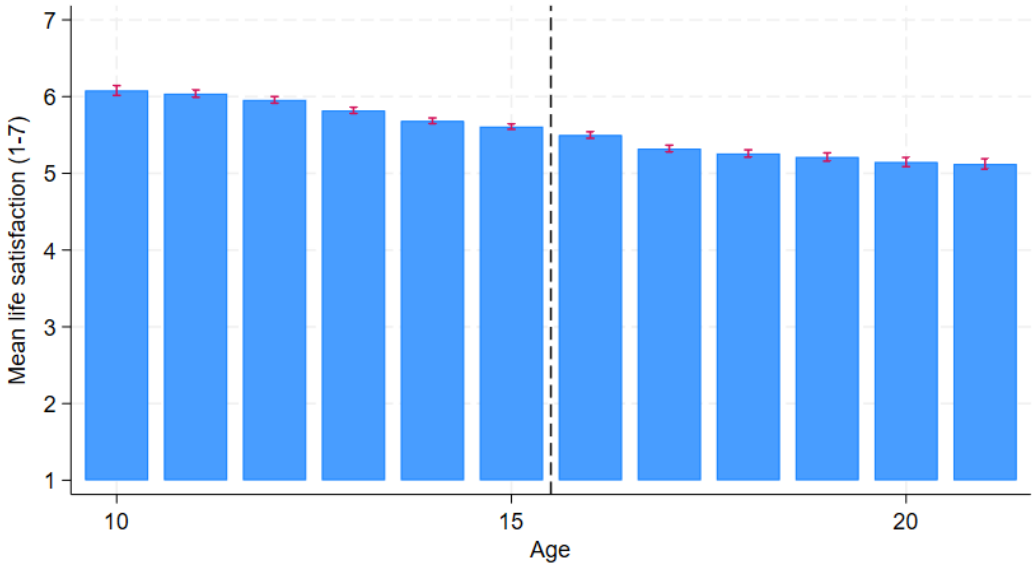
At the reading age of 10[3], children are considered capable of responding to the ONS-4 personal wellbeing questions by the Child Outcomes Research Consortium (CORC). Furthermore, beyond age 11, responses to questionnaire modules such as the Strengths and Difficulties Questionnaire and the Short Moods and Feelings Questionnaire can also be used. Psychometric testing shows that these responses are stable, valid, and reliable (Goodman et al., 1998; Goodman, 2001).

The UK Household Longitudinal Study "Understanding Society" includes a variant of an evaluative life satisfaction question for individuals between 10 and 15 in their youth panel, asking "How do you feel about life as a whole? 1: Not at all happy, 7: Completely happy". This question blends the evaluative notion of 'life as a whole' with the experiential wellbeing concept 'happiness'; it is not clear *a priori* how to interpret individual's responses. We investigate this empirically, as many of the individuals (85%) responding to this question in the youth panel go on to respond to the standard ONS life satisfaction question in the adult

[3] Reading age and chronological age do not map one-to-one. In 2022-23 73% of 10-11-year-olds were at the expected reading age in the UK (DfE, 2024).

panel upon turning 16. Figure 1 shows the mean response to the life evaluation question in Understanding Society from ages 10-21. There is no clear discontinuity in mean response over the cut-off when the phrasing of the question changes both visually and when evaluating the difference in means with a series of t-tests. These results hold true for the full sample and for a series of demographic subgroups, results are reported in Appendix Table A1. The response distributions also show minimal variation over the cutoff, reported in Appendix Figure A1. This provides evidence that the two questions are answered in the same way by those under and over 16 years of age. Therefore, we recommend that this question can be considered equivalent to the standard ONS life satisfaction question[4].

Figure 1. Mean Life Evaluation by Age



Notes: Blue bars represent mean life satisfaction over age group. Red lines indicate 95% confidence intervals of the mean.

The Active Lives Children and Young People Survey (ALS) directly asks children between 11 and 16 the standard ONS life satisfaction question[5]. It is therefore possible to run direct tests of internal and external validity, common in the measurement validation literature. At each age, in both datasets, life satisfaction follows a distribution with a mode of 7-8/10 and a distinct left skew – the same as that which is seen for adults. Responses also show congruent validity, correlating accordingly with similar concepts such as trust, eudaimonia, happiness, and loneliness. See Appendix Table A2 and Figure A2.

It seems clear from both sets of results that children above the age of 10 are suitably equipped to evaluate their lives 'on the whole'. It is also clear that the Understanding Society youth question could be considered an evaluative judgement and can be used

[4] This assumption is not uncommon in wellbeing research. Small variations in question wording across surveys are common and are considered acceptable where the underlying concept of evaluative wellbeing is clear.
[5] *"How satisfied are you with your life nowadays? 0: Not at all, 10: Completely"*.

interchangeably with the usual life satisfaction question. We therefore recommend that life satisfaction measures and WELLBYs are used in cost-benefit analyses concerning children aged 10 and over. On the grounds of equity, it is further recommended that these are valued at the same price as the current WELLBY for adults; £15,920 in 2024 prices (HMT, 2021). These can be included in addition to the monetised future beneficial outcomes from an improved childhood without fear of double-counting, as they relate only to wellbeing in the current period.

3. Below Age 10

Under the premise that children below 10 cannot, on aggregate, consistently assess their own life satisfaction, the use of an evaluative subjective wellbeing measure becomes conceptually fraught. If wellbeing lies in how we evaluate our own lives to be going (Dolan et al., 2008), a child below 10 ceases to be a welfare subject. Implicitly, they cannot have wellbeing. This is currently how policy around young children's wellbeing operates, wellbeing can only be valued once a person can report it for themselves. This is clearly unjust. It is uncontroversial to suggest that young children's wellbeing should be valued just as much as adult's wellbeing. Solutions to this problem fall into three categories.

1. Change theory of wellbeing, and hence metric, to one that everyone can experience and answer.

The main candidate solution for this would be hedonic wellbeing, using a measure such as self-reported day-to-day happiness. Children at ages below 10 show superior reliability in responding to questions about their happiness 'in the moment' than they do for evaluative questions (Fallon & Schwab-Stone, 1994). In fact, children from ages 5-7 are already asked a happiness question in the Active Lives Survey[6]. The benefits of the happiness approach are that there could be a consistent unit of measurement across the life course[7]. However, there are multiple issues with this suggestion. Initially, the cost of collecting repeated experiential happiness data is currently prohibitively high, where it relies on methods such as the Day Reconstruction Method (DRM) (Kahneman, 2004) or Experience Sampling Method (ESM) (Hektner, 2007). Secondly, it would require overhauling the current WELLBY method for adults in favour of experienced wellbeing. In the long term this could be beneficial, however it would bring wellbeing policy research 'back to square one' in the short term. Finally, young children have not yet been shown to consistently respond to questions about happiness or other affective states below age 5. The use of happiness as a universal indicator seems akin to pushing this problem a few years down the road, rather than solving it[8].

[6] Although, this is heavily simplified question and would not be suitable for analysis using common wellbeing approaches. Children are shown a smiling face, a straight face, and sad face and asked to choose which one they feel like.

[7] This is welfare invariablism, see Lin (2018).

[8] There may be some capacity for parent or teacher reported happiness at ages younger than 5, but there is currently no evidence on the reliability of other-reported happiness that we are aware of.

Another possible candidate for a universal all-age measure of wellbeing is an objective list, such as the OECD's Better Life Index. Using an objective list could seemingly allow adults and children to be assessed on the same criteria, as it could be completed on behalf of a child more easily. However, in practice, what makes life good for children may not be the same as for adults. Although this could be solved if the objective list were different for children and adults, this would both defeat the purpose of a universal measure. Furthermore, the benefits of capturing subjective wellbeing would be lost, as individuals could not decide for themselves how their lives are deemed to be going. It seems there is no one measure which could adequately capture subjective wellbeing across the life course.

2. Allow different measures of subjective wellbeing for different ages.

Perhaps, day-to-day happiness for those under 10 and life satisfaction for those 10 and over. This would be a simple approach which allows the main WELLBY framework to remain intact, if a suitable way to estimate a monetary value for happiness scores is found. However, for those under 5, the problem that these questions are not reliably answerable remains. There may be some capacity for parent- or teacher- reported happiness at ages younger than 5, but there is currently no evidence on the validity of other-reported happiness that we are aware of. In the current framework, it would be necessary to translate these self- or other-reported happiness scores for children ages 2-9 into a life satisfaction score which can then be monetised in terms of WELLBYs. One could simply equate one point on a momentary happiness scale with one point on a life satisfaction scale, although this is likely to be an oversimplification, given that these questions capture distinct concepts.

Another solution would be to measure the effect that a one-point change in momentary happiness has on life satisfaction, using this to derive the relevant WELLBYs. This leads to a final, related possibility for monetising children's wellbeing effectively.

3. Estimate what a young child would have scored on the life satisfaction scale if they could answer it.

The most pragmatic response might be to accept that life satisfaction cannot be directly measured for young people and instead to proxy for it using the best available evidence. By estimating the relationship between a different measure of subjective wellbeing and life satisfaction for this group, it would be possible to approximate a hypothetical life satisfaction score. However, if young children are incapable of making evaluative life judgements, how can we analyse this relationship in the absence of a life satisfaction outcome? The closest comparison group from which this could be estimated would be the closest age group that is considered able to evaluate their lives; 10-15-year-olds.

There are several measures that could be used to estimate this hypothetical life satisfaction score. Self-reported momentary happiness[9] would be likely to capture a large amount of this

[9] From age 5 and above.

variation in wellbeing. Using data from the Active Lives Survey, a one-point increase in happiness yesterday (0-10 scale) is associated with a 0.63 increase in life satisfaction (0-10 scale) for 11-16-year-olds. This translation could be directly used on young children's momentary happiness scores to estimate a C-WELLBY.

Another possibility would be to use a pre-existing child specific wellbeing measure. For this estimation to be useful, it would need to be a measure of wellbeing which is both commonly used in policy evaluation and validated for children both in the target age group (2-9) and in the test age group (10-16). An obvious choice in this case is the Strengths and Difficulties Questionnaire (SDQ), which is commonly collected for these ages and is highly validated (Goodman, 1997). Below age 11, either parent- or teacher- reports are recommended to measure the SDQ, overcoming issues around young children's inability to self-report. Parent and teacher reports of SDQ have been shown to correlate reasonably well with children's self-reports at ages above 10, both in the literature (Goodman et al., 1998; Goodman et al., 2010) and in our own analyses[10]. More recent research using the British Child and Adolescent Mental Health Survey shows that the distribution of self-reported SDQ scores is most similar to parent-reported SDQ, rather than teacher-reported SDQ (Ford & Parker, 2016).

The SDQ comprises 20 questions split into four groups: emotional symptoms, conduct problems, hyperactivity, and peer relationship problems. Each question is scored on a 0-2 scale where 0 is no difficulties and 2 is the most difficulty. The SDQ is generally reported as a total difficulties score (all questions summed), or as an internalising score (SDQ-I) which includes only emotional symptoms and peer relation scores. The SDQ hence ranges 0-40 while the SDQ-I ranges 0-20. It is unclear which sections of the SDQ will be the strongest predictors of life satisfaction. The emotional sub-score alone is the most conceptually similar to life satisfaction, however this is rarely reported independently in intervention evaluations and so is not assessed individually in this report.

The next section provides an example of the use of mapping functions to predict WELLBYs using momentary happiness, SDQ, and SDQ-I scores. It attempts to answer two questions:

- How accurately can you predict a WELLBY using the SDQ and momentary happiness scores?
- How feasible is it to extrapolate this figure from the sample group (11–16-year-olds) to the target group (2–9-year-olds)?

4. Mapping Happiness and SDQ responses to life satisfaction

By analysing the relationship between two measures in a crosswalk dataset containing both measures, it is possible to produce a function which maps a reported measure onto a target measure. This technique draws on the idea of 'mapping functions', which are a NICE[11]

[10] We investigate the correlations between self, parent, and teacher reports using the Millenium Cohort Study, results are reported in Appendix Table A3

[11] National Institute for Health and Care Excellence.

recommended solution to a similar problem in health economics, where a given outcome (usually QALYs) has not been measured directly and so must be estimated using other health outcomes (Mukuria et al., 2024). There have been over 200 published mapping studies since 2007, discussed in a meta-analysis by Mukuria et al. (2019)[12]. Using insights from this expanding literature and drawing on the ISPOR[13] Task Force Report for good practices in mapping studies (Wailoo et al., 2017), I estimate mapping functions which translate self-reported SDQ and Happiness scores into life satisfaction responses (WELLBYs). I further investigate how these mapping functions vary with age, providing evidence on the feasibility of extrapolating to lower age groups. If this relationship holds true at ages below 10, this translation can be applied in policy analysis to estimate changes in life satisfaction for children aged 2-9. Descriptive statistics by age for life satisfaction, SDQ, SDQ-I, and happiness are reported in Table 1; distributions are reported in Appendix Figures A1, A2, B1, B2.

Table 1. Descriptive Statistics

			Age				
Dataset	**Measure**	**Average**	**11**	**12**	**13**	**14**	**15**
UKHLS	**Life Satisfaction**	7.90	8.17	8.04	7.81	7.62	7.55
	(0-10)	(1.88)	(1.78)	(1.82)	(1.88)	(1.91)	(1.91)
	SDQ	10.70	10.38	10.44	10.71	10.83	11.07
	(0-40)	(5.77)	(5.81)	(5.74)	(5.70)	(5.74)	(5.68)
	SDQ-I	4.67	4.52	4.41	4.64	4.76	5.01
	(0-20)	(3.32)	(3.21)	(3.22)	(3.37)	(3.39)	(3.37)
	N	21,027	3,468	3,614	3,528	3,656	3,412
ALS	**Life Satisfaction**	6.50	7.16	6.94	6.54	6.22	5.94
	(0-10)	(2.66)	(2.60)	(2.53)	(2.61)	(2.65)	(2.68)
	Happiness	6.87	7.11	6.74	6.39	6.14	5.93
	(0-10)	(2.76)	(2.66)	(2.60)	(2.67)	(2.71)	(2.75)
	N	337,657	72,528	81,290	76,347	66,621	56,440

Notes: UKHLS data is from UK Household Longitudinal Study Youth panel 2009-2019, where SDQ and SDQ-I are reported bi-annually. ALS data is from Active Lives of Children and Young People Survey with six annual cross-sections from 2017-2023.

Modelling

After testing multiple model specifications (linear, quadratic, cubic; with/without controls; with/without scores disaggregated) and model types (robust OLS, Tobit, fixed effects OLS), both within-sample and using tenfold cross-validation - a robust linear OLS model with controls for age, gender, and year was chosen. Standard errors are clustered at the individual level. Full details of the model selection process are discussed in Appendix B. Reported in

[12] See Dakin et al. (2023) for an extensive database of mapping studies from 2013-2023 from the Oxford Health Economics Research Centre (HERC).

[13] The International Society for Pharmacoeconomics and Outcomes Research.

Table 2 are the mapping coefficients for both happiness, SDQ, and SDQ-I, shown as an overall figure and split by age, from the specification:

$$LifeSatisfaction_i = \alpha_1 + \theta_1 W_i + X_i\beta + \varepsilon_i \quad (1)$$

Where X_i is a vector of controls: age (one-year bins), gender (binary), and year (one-year bins; where appropriate). W_i is the reported outcome measure and θ_1 is the 'mapping coefficient' from the linear model.

Results

There is little variation in mapping over different age groups, with the difference between the average coefficient and the youngest age group's coefficient only 10% of the average for each measure. There is a rough trend as age decreases which is downward for happiness and upward for SDQ and SDQ-I. This trend is not significant enough in either case to confidently assume that it continues beyond 11 and extrapolate accordingly. Instead, it is recommended to use the youngest age-specific coefficient available for translations (SDQ: -0.146[14], SDQ-I: -0.246, Happiness: 0.546).

Table 2. Mapping Coefficients for SDQ, SDQ-I, and Happiness by Age.

		Age				
Measure	**Average**	**11**	**12**	**13**	**14**	**15**
SDQ θ_1 *(0-40)*	-0.166*** [-0.171, -0.161]	-0.146*** [-0.156, -0.135]	-0.160*** [-0.170, -0.149]	-0.175*** [-0.186, -0.164]	-0.175*** [-0.186, -0.164]	-0.170*** [-0.181, -0.158]
SDQ-I θ_1 *(0-20)*	-0.279*** [-0.289, -0.270]	-0.246*** [-0.266, -0.227]	-0.274*** [-0.293, -0.254]	-0.286*** [-0.306, -0.266]	-0.293*** [-0.313, -0.274]	-0.288*** [-0.309, -0.268]
Happiness θ_1 *(0-10)*	0.603*** [0.600, 0.606]	0.546*** [0.536, 0.556]	0.575*** [0.568, 0.581]	0.613*** [0.606, 0.620]	0.621*** [0.614, 0.628]	0.628*** [0.621, 0.636]

Notes: Figures are regression estimates from equation (1). SDQ analysis uses data from Understanding society youth panel. Happiness analysis uses data from the Active Lives Children and Young People Survey. 95% confidence intervals are in square brackets. ***p<0.001, **p<0.01, *p<0.05

A statistically significant relationship is a necessary but not sufficient condition for a successful mapping of two measures. Table 3 shows the predictive power of each of the

[14] Understanding Society also asks the life satisfaction to 10-year-olds, although it is not clear that mapping will be reliable as the self-reported SDQ is not validated for this age group. We do also calculate a mapping coefficient for this group and find that it is the same as for 11-year-olds (-0.147), lending some credence to the reliability of extrapolating this figure downwards.

mappings. At the tail ends of the SDQ distribution there are weaknesses; it is not possible to capture life satisfaction levels below 3.7 (5% of the sample are below this level) and the upper bound is beyond the maximum scale point (>10). For the SDQ-I, the tail ends of the distribution show similar issues. It is not possible to capture levels below 3.7, or above 9.6 (above which 35% of the responses sit). Happiness shows superior distributional qualities, with a larger range, encompassing 78% of the life satisfaction responses[15]. For all measures, the predicted standard deviation is roughly halved, this 'shrinking' is simply an artefact of regression-based mapping, as all unexplained idiosyncratic variation is stripped away as noise. Figures B1-6 in Appendix B graph the predicted and actual mean values of each measure at each level of life satisfaction and show the predicted and actual cumulative distribution functions for each measure. For ease of use, only linear coefficients are reported in this paper, quadratic functions show superior distributional properties (see Appendix B1-6) and should be further investigated in future.

Conceptual Comparability

Beyond statistical relationships, it is necessary to consider the conceptual comparability between the measures in question. Although day-to-day happiness is an important constituent of life satisfaction (Layard & De Neve, 2023), it is only one aspect of a wider life evaluation, so solely using this might miss important information. The full SDQ captures a broader range of items, split into four categories: emotional symptoms, conduct problems, hyperactivity, and peer relationship problems, while the SDQ-I considers only the emotional and conduct problems subsections. The full SDQ criteria encompass both internal and external aspects of life, which deviates from the subjective wellbeing account conceptually, although may still work as a strong predictor of life satisfaction. The SDQ-I criteria are the most well aligned with a life satisfaction evaluation, capturing evaluations of a series of mental states.

[15] Interventions for young people may focus on specific groups with low prior wellbeing, in this instance being unable to capture the lower end of the distribution may be problematic. We test this empirically, both using quantile regressions and by re-weighting our representative UK sample to match a theoretical low wellbeing group. In each instance the low wellbeing group shows a roughly 25% larger absolute mapping coefficient. This makes the reported mapping functions lower bounds for low wellbeing groups and suggests future mapping would benefit from lookup tables with functional forms where prior group wellbeing can be accounted for.

Table 3. Actual and Predicted Life Satisfaction

Statistic	Understanding Society			Active Lives Survey	
	Life satisfaction	Predicted life satisfaction (SDQ)	Predicted life satisfaction (SDQ-I)	Life satisfaction	Predicted life satisfaction (Happiness)
Mean	7.90	7.90	7.90	6.60	6.60
(SD)	(1.88)	(0.99)	(0.95)	(2.59)	(1.67)
Min	0.29	3.73	3.65	0	2.08
Max	9.71	10.18	9.61	10	9.29
Range	9.42	6.45	5.96	10	7.21

Notes: Predictions are based on the estimated $\widehat{LifeSatisfaction_i}$ from equation (1) ($LifeSatisfaction_i = \alpha_1 + \theta_1 W_i + X_i\beta + \varepsilon_i$). Predictions are based on the full sample (ages 11-16).

5. Discussion

Recommendations

There is no consensus on the measurement of children's wellbeing in policy. Accordingly, there is a lack of consistent measurement of children's wellbeing across statistical agencies and organisations. For children aged 10 and over, there is strong evidence that the ONS's life satisfaction question can be adequately understood and answered[16]. We recommend the use of this standard phrasing for 10-16-year-olds to improve comparability with adult's wellbeing. In essence, as soon as a young person can answer the life satisfaction question, they should be treated no differently to adults. We recommend that this is valued at the same £15,920 as the adult WELLBY.

For children aged under 10, there is a need to translate more accessible measures of wellbeing into C-WELLBYs for their appropriate inclusion in cost-benefit analyses. The coefficients reported here introduce the use of mapping functions to wellbeing analysis for children, using common techniques from the health economic literature (Wailoo et al., 2017). We recommend the use of the SDQ Total Difficulties Score translated to WELLBYs using a linear coefficient of -0.146, for the age group 2-9. Given its superior range, we recommend the use of parent reported SDQ to undertake this mapping where possible, while if this is not available, teacher reported SDQ is also acceptable[17]. If the SDQ is not available, we recommend the use of happiness as a quick, easy way to estimate WELLBYs for children, using a linear coefficient of 0.546. Where confidence intervals are to be included around parameter estimates, we recommend the use of those reported in Table 2 (SDQ: [-

[16] The Good Childhood Report 2013 (Rees et al., 2013) recommends that the life satisfaction question can be answered from age 8 and above, however we only have data to support this claim from age 10.

[17] It should be noted that this recommendation relies on the assumption that the self-reported SDQ maps sufficiently closely to parent- and teacher-reported SDQ.

0.156, -0.135]; Happiness: [0.536, 0.556]). We recommend that this C-WELLBY is also valued at £15,920.

Limitations and Extensions

There remains several limitations and extensions to this approach. As discussed, mapping of this sort can show improved predictive accuracy using more complex functional forms. The use of quadratic or cubic functions allows for differential effects based on where in the distribution respondents begin, while the use of interactions with demographics, such as age and gender, allow differential demographic mapping effects. For the purposes of this introductory paper, these functional forms were not reported. While predictive accuracy can be improved, the degree to which it improves was not deemed sufficiently substantive to overcome the additional complexity. Future research would benefit from the production of lookup tables, which allow ease of translation with more complex functional forms.

Furthermore, a mapping where the predictor only captures one aspect of the outcome will systematically underestimate the corresponding change following an intervention. While both SDQ (mental ill health) and happiness (experiential wellbeing) explain a large proportion of the observed variance in life satisfaction, they do not capture all possible sources of variation that may arise from an intervention. It is therefore necessary to treat mapping functions as a lower bound estimate. In the case of the C-WELLBY, where a direct WELLBY measurement cannot be made, the use of a purely 'mapped' C-WELLBY will introduce a systematic downward bias on the value of wellbeing for young children. This would be unfair, and practical solutions to this should be investigated.

Further research would also benefit from more suitable crosswalk datasets. The mapping of SDQ currently relies on a linear translation of life satisfaction from a 1-7 scale to a 0-10 scale, as this is how the question is formulated in Understanding Society. More accurate functions could be estimated with a full 0-10 scale introducing more variance. Also, beyond quantitative analyses, a deeper understanding of children's perceptions of their lives and what makes them go well would give further insight into the feasibility of proxying for life satisfaction.

References

Dakin, H., Abel, L., Burns, R., Koleva-Kolarova, R., & Yang, Y. (2023). HERC database of mapping studies, Version 9.0 (Last updated: 1st December 2023). https://doi.org/10.5287/ora-avomzqyvb. Available at: http://www.herc.ox.ac.uk/downloads/herc-database-of-mapping-studies.

Department for Education. (2024). School results for 10- to 11-year-olds. In Ethnicity facts and figures. GOV.UK. Retrieved [12/02/25], from https://www.ethnicity-facts-figures.service.gov.uk/education-skills-and-training/7-to-11-years-old/school-results-for-10-to-11-year-olds/latest

Dolan, P., Peasgood, T., & White, M. (2008). Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *Journal of economic psychology*, 29(1), 94-122.

Fallon Jr, T., & Schwab-Stone, M. (1994). Determinants of reliability in psychiatric surveys of children aged 6–12. *Journal of Child Psychology and Psychiatry*, *35*(8), 1391-1408.

Ford, T., & Parker, C. (2016). Emotional and behavioural difficulties and mental (ill) health. *Emotional and behavioural difficulties*, *21*(1), 1-7.

Frijters, P., & Krekel, C. (2021). *A handbook for wellbeing policy-making: History, theory, measurement, implementation, and examples* (p. 454). Oxford University Press.

Goodman, A., Lamping, D. L., & Ploubidis, G. B. (2010). When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): data from British parents, teachers and children. *Journal of abnormal child psychology*, *38*, 1179-1191.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of child psychology and psychiatry*, *38*(5), 581-586.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(11), 1337-1345.

Goodman, R., Meltzer, H., & Bailey, V. (1998). The Strengths and Difficulties Questionnaire: A pilot study on the validity of the self-report version. *European child & adolescent psychiatry*, *7*(3), 125-130.

Hektner, J. M. (2007). *Experience sampling method: Measuring the quality of everyday life*. Sage Publications.

HM Treasury. (2021). Wellbeing Guidance for Appraisal: Supplementary Green Book Guidance. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1005388/Wellbeing_guidance_for_appraisal_-_supplementary_Green_Book_guidance.pdf

Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. *Science*, *306*(5702), 1776-1780.

Layard, R., & De Neve, J. E. (2023). *Wellbeing*. Cambridge University Press.

Lin, E. (2018). Welfare invariabilism. *Ethics*, *128*(2), 320-345.

Little, A., & Parkes, I. (2024). Child Wellbeing-Years (C-WELLBYs): Valuing children's wellbeing in economic and social value analysis. State of Life Report. Available at: https://drive.google.com/file/d/1_hwHw4_j4M6CybdpMihNMtQulKAvIfeQ/view

Mukuria, C., Franklin, M., & Hinde, S. (2024). Mapping functions for the PHQ-9 and GAD-7 to generate EQ-5D-3L for economic evaluation. The European Journal of Health Economics, 1-8.

Mukuria, C., Rowen, D., Harnan, S., Rawdin, A., Wong, R., Ara, R., & Brazier, J. (2019). An updated systematic review of studies mapping (or cross-walking) measures of health-related quality of life to generic preference-based measures to generate utility values. Applied health economics and health policy, 17, 295-313.

OECD. (2019). *The OECD Better Life Initiative: Measuring Well-being and Progress.* OECD Publishing.

Rees, G., Goswami, H., Pople, L., Bradshaw, J. R., Keung, A., & Main, G. (2013). *The Good Childhood Report 2013.* The Children's Society.

Sport England. (2023). *Active Lives Children and Young People Survey, 2017–2023*. Sport England.

University of Essex, Institute for Social and Economic Research. (2023*). Understanding Society: Waves 1–12, 2009–2022* (15th ed.). UK Data Service. SN: 6614. https://doi.org/10.5255/UKDA-SN-6614-15

Wailoo, A. J., Hernandez-Alava, M., Manca, A., Mejia, A., Ray, J., Crawford, B., ... & Busschbach, J. (2017). Mapping to estimate health-state utility from non–preference-based outcome measures: an ISPOR good practices for outcomes research task force report. Value in Health, 20(1), 18-27.

APPENDIX A

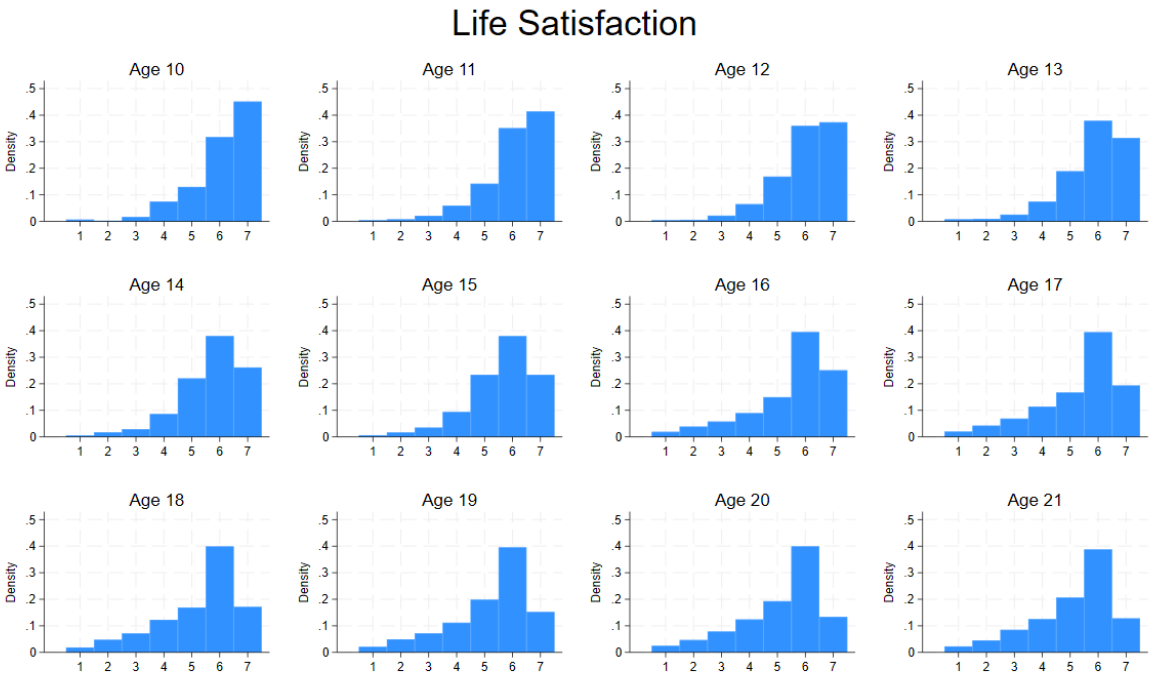
Table A1. Difference in means between life satisfaction responses at each age group.

	Full sample		Gender				Ethnicity							
			Male		Female		White		Mixed		Asian		Black	
Age														
10	-0.04	(0.317)	-0.035	(0.557)	-0.045	(0.404)	-0.053	(0.266)	0.168	(0.306)	-0.013	(0.931)	-0.038	(0.789)
11	-0.077	(0.018)	-0.009	(0.854)	-0.139	(0.002)	-0.077	(0.047)	0.004	(0.973)	-0.187	(0.094)	-0.057	(0.626)
12	-0.139	(0.000)	-0.056	(0.182)	-0.216	(0.000)	-0.12	(0.001)	-0.232	(0.034)	-0.057	(0.574)	-0.265	(0.035)
13	-0.136	(0.000)	-0.039	(0.299)	-0.212	(0.000)	-0.134	(0.000)	-0.231	(0.040)	-0.123	(0.182)	-0.044	(0.701)
14	-0.077	(0.003)	-0.099	(0.005)	-0.064	(0.090)	-0.069	(0.024)	-0.149	(0.171)	-0.083	(0.313)	-0.145	(0.168)
15	-0.108	(0.000)	-0.118	(0.003)	-0.097	(0.013)	-0.098	(0.003)	0.004	(0.974)	-0.175	(0.042)	-0.092	(0.420)
16	-0.176	(0.000)	-0.177	(0.000)	-0.174	(0.000)	-0.186	(0.000)	-0.184	(0.148)	-0.177	(0.063)	-0.144	(0.232)
17	-0.064	(0.053)	-0.079	(0.096)	-0.051	(0.264)	-0.076	(0.051)	-0.058	(0.681)	0.024	(0.809)	-0.123	(0.349)
18	-0.046	(0.210)	-0.107	(0.042)	0.011	(0.833)	-0.036	(0.400)	0.068	(0.653)	-0.217	(0.065)	0.112	(0.434)
19	-0.066	(0.110)	-0.163	(0.009)	0.014	(0.797)	-0.085	(0.079)	0.003	(0.986)	0.008	(0.952)	-0.16	(0.313)
20	-0.025	(0.593)	-0.008	(0.914)	-0.038	(0.528)	-0.032	(0.555)	-0.218	(0.212)	0.091	(0.576)	0.081	(0.641)
21	-0.029	(0.565)	-0.095	(0.227)	0.021	(0.753)	-0.019	(0.753)	-0.076	(0.721)	-0.173	(0.355)	-0.169	(0.372)
22	0.015	(0.792)	0.011	(0.902)	0.016	(0.826)	0.023	(0.729)	0.281	(0.215)	0.016	(0.935)	-0.005	(0.980)
23	-0.019	(0.768)	-0.013	(0.904)	-0.027	(0.745)	0.004	(0.954)	-0.136	(0.559)	-0.075	(0.755)	-0.075	(0.765)
24	-0.066	(0.405)	-0.16	(0.212)	-0.009	(0.931)	-0.074	(0.419)	0.054	(0.845)	-0.14	(0.655)	-0.034	(0.910)

Notes: Age 10 indicates a difference in mean between life satisfaction response for the sample of 10-year-olds and 11-year-olds, age 11 is the difference between age 11 and age 12 etc. A more statistically significant difference is shown by darker green shading.

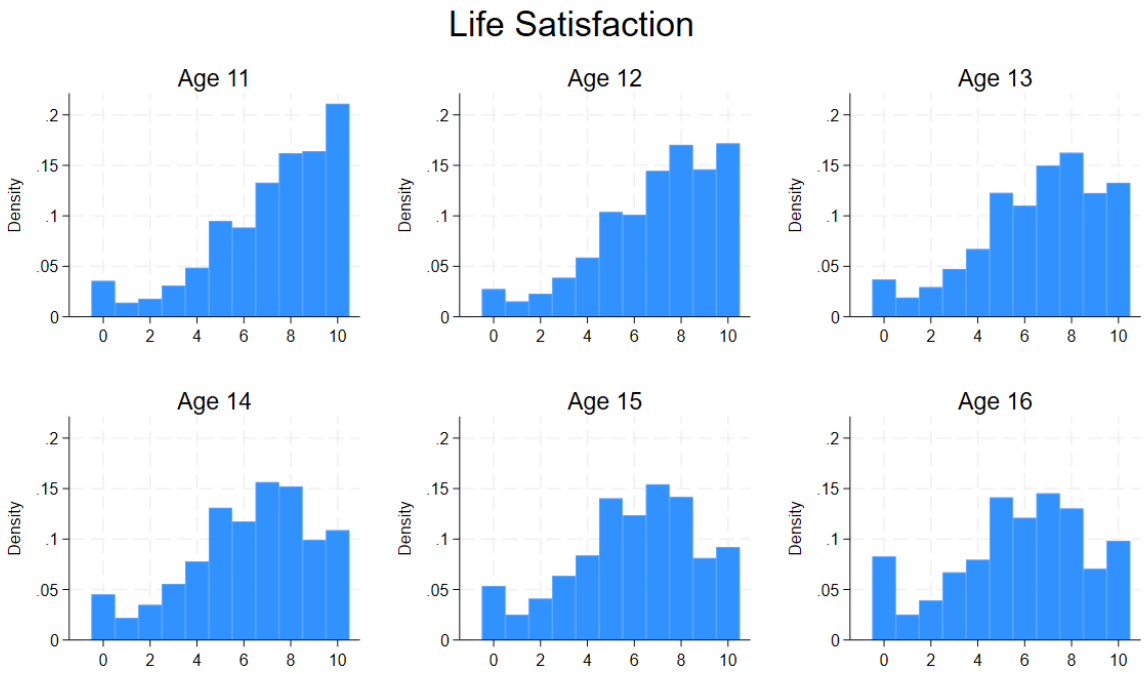
There is no discontinuity in mean response between ages 15 and 16, where the wording of the question changes. Although there is a significant difference in mean response over the threshold, this is in keeping with the observed prior pattern, where life satisfaction is decreasing from ages 11 to 18

Figure A1: Histograms showing life satisfaction distribution at each age group.



Notes: Data is from Understanding Society.

Figure A2. Histograms showing life satisfaction distribution at each age group



Notes: Data is from Active Lives of Children and Young People Survey.

Table A2. Correlation Table for Life Satisfaction and Related Measures

	Happiness	**Feelings of worthwhileness**	**Trust in people**	**Loneliness**
Life Satisfaction	0.644	0.748	0.348	-0.509
	(0.000)	(0.000)	(0.000)	(0.000)

Notes: Pairwise correlations are reported. Data is from Active Lives of Children and Young People Survey. Life Satisfaction is measured by the question "How satisfied are you with your life nowadays? 0-10". Happiness is measured by the question "How happy did you feel yesterday? 0-10". Feelings of worthwhileness is measured by "To what extent are the things you do in life worthwhile? 0-10". Trust in people is measured by "How much do you feel you can trust people who are a similar age to you? 1-4". Loneliness is measured by "How often do you feel lonely? 1-5".

Table A3. Correlation Matrix for Child-, Parent-, and Teacher-Reported SDQ and SDQ-I

SDQ

	Child	**Parent**	**Teacher**
Child	x		
Parent	0.44 (0.00)	x	
Teacher	0.46 (0.00)	0.53 (0.00)	x

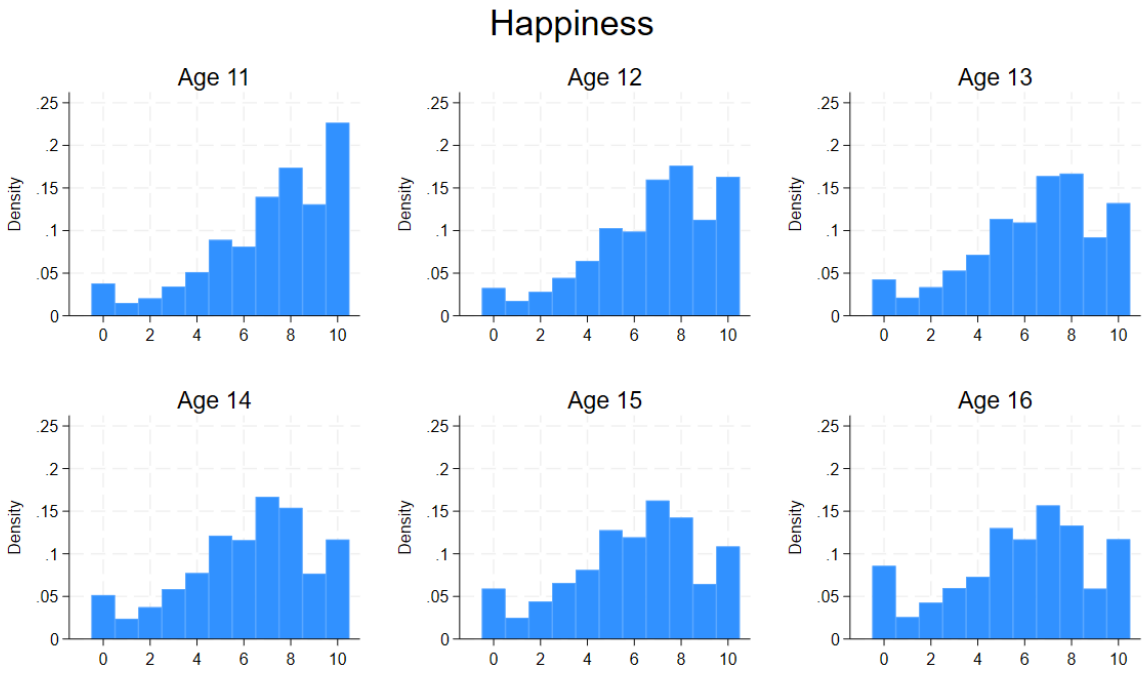
SDQ-I

	Child	**Parent**	**Teacher**
Child	x		
Parent	0.45 (0.00)	x	
Teacher	0.32 (0.00)	0.44 (0.00)	x

Notes: Pairwise correlations are reported. Data is from Millenium Cohort Study. Child-Teacher correlation uses data from wave 4 (age 7), Parent-Teacher correlation uses data from wave 6 (age 14), Parent-Child correlation uses data from wave 7 (age 17).

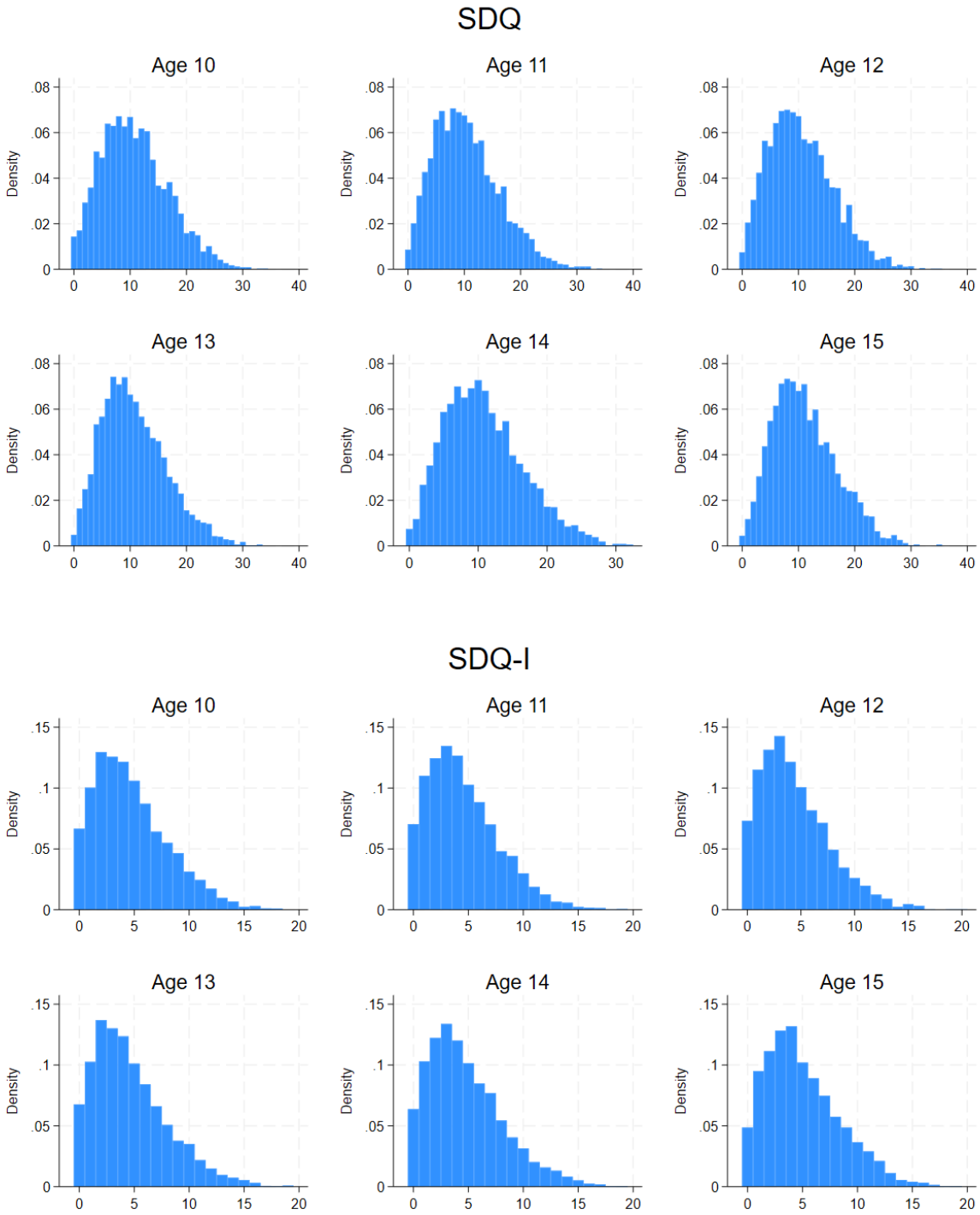
APPENDIX B

Figure B1. Histograms showing happiness distribution at each age group



Notes: Data is from Active Lives of Children and Young People Survey.

Figure B2. Histograms showing SDQ and SDQ-I distribution at each age group



Notes: Data is from Understanding Society Youth Panel.

Information on Mapping

Data

Data was drawn from two country-wide representative studies: the UK Household Longitudinal Study youth panel (UKHLS-Y) and Active Lives of Children and Young People Survey (ALS).

UKHLS, also known as Understanding Society, is an extensive panel survey that tracks economic, social, and health conditions of around 40,000 households across the United Kingdom. The self-competed survey employs a stratified sampling design, ensuring representativeness across various population subgroups, and covers the years 2009-2021. The youth panel comprises respondents solely age 10 to 15. UKHLS-Y contains annual data on life satisfaction and biennial data on SDQ and SDQ-I. The life satisfaction question is scored on a 1-7 scale, which is translated to a 0-10 scale prior to analysis[18].

The Active Lives of Children and Young People Survey is a repeated cross-section that tracks the health and wellbeing of around 110,000 young people in the UK each year. It is self-completed and employs a stratified sampling design which ensures a representative population. The ALS contains annual data on life satisfaction and happiness scored on a 0-10 scale for the years 2017-2023.

Modelling

Mapping functions use a predictive regression model to estimate the relationship between a reported outcome and a target measure. In keeping with ISPOR recommendations for mapping studies, multiple model specifications are analysed, including using total outcome scores, question-level scores, and response category scores as binary indicators, both linearly and with 2nd and 3rd order polynomials. All models were also tested with and without controls for age, gender, and year. Model fit was assessed using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). AIC and BIC are penalised likelihood statistics where AIC measures the relative amount of information lost by representing a given data generating process with the model in question. BIC is similar but has a stronger penalty for the inclusion of additional parameters and so is more likely to favour simpler models.

Table B1. Mapping Function Specifications

(1)	$LifeSatisfaction_i = \alpha_1 + \theta_1 W_i + X_i\beta + \varepsilon_i$
(2)	$LifeSatisfaction_i = \alpha_1 + \theta_1 W_i + \theta_2 {W_i}^2 + X_i\beta + \varepsilon_i$
(3)	$LifeSatisfaction_i = \alpha_1 + \theta_1 W_i + \theta_2 {W_i}^2 + \theta_3 {W_i}^3 + X_i\beta + \varepsilon_i$

[18] Using the formula recommended in Frijters and Krekel (2021). This linear conversion uses the equidistance assumption and the maximum information principle. It assumes cardinality. There is some potential for error to be introduced here, but it is unlikely to substantively affect the final mapping function.

(4) $$LifeSatisfaction_i = \alpha_1 + \sum_{j=1}^{N} \theta_j W_{i\ question\ j} + X_i\beta + \varepsilon_i$$

(5) $$LifeSatisfaction_i = \alpha_1 + \sum_{j=1}^{N} \theta_j W_{i\ question\ j} + \sum_{j=1}^{M} \theta_{j+N} W_{i\ question\ j}{}^2 + X_i\beta +$$

(6) $$LifeSatisfaction_i = \alpha_1 + \sum_{j=1}^{N} \theta_j W_{i\ question\ j} + \sum_{j=1}^{N} \theta_{j+N} W_{i\ question\ j}{}^2 + \sum_{j=1}^{N} \theta_{j+2N} W_{i\ question\ j}{}^3 + X_i\beta + \varepsilon_i$$

(7) $$LifeSatisfaction_i = \alpha_1 + \sum_{j=1}^{N} \theta_j W_{i\ response\ category\ j} + X_i\beta + \varepsilon_i$$

All models were repeated without controls for age, gender, and year.

In the absence of a universal best measure of model prediction in mapping studies, mean absolute error (MAE), mean error (ME), and root mean squared error (RMSE) were all considered. Initially, prediction was assessed using full sample error statistics. To ensure internal validity, tenfold cross-validation was then used to assess out of sample prediction, as is recommended by ISPOR. Both robust OLS and Tobit models were tested for model fit using tenfold cross-validation, given evidence of non-normality and floor/ceiling effects in the measures used[19]. Table B2 reports descriptive statistics, including skewness, kurtosis, and presence of floor/ceiling effects for each measure. For data with a panel structure, standard errors were clustered at the level of the individual. To select a preferred model, there is no mechanistic rule for mapping, instead a researcher must use specific judgement to assess a model's suitability for the functions intended application. All model fit results are reported in Table B3.

In all cases, for all measures, inclusion of covariates improved model fit, justifying the addition of controls for age, gender, and year empirically. In almost all cases, the use of Tobit regression models did not improve model fit over robust OLS using cross-validation. Robust OLS was selected as the preferred model. Some improvement in model fit was observed when using more complex mapping functions, however improvements tended to be minimal[20]. For the sake of parsimony, only linear, total score coefficients are reported as main results. There are no instances within a standard response range where using different models leads to

[19] Models using fixed effects were considered but were not included. A fixed effect regression will mechanistically produce a lower coefficient, as it relies only on within-person variation. Where some crosswalk datasets are panels and some are not, using fixed effects only for those panel crosswalks will downwardly bias the mapping of these measures in comparison to the others. For consistency it was decided to exclude fixed effects entirely.

[20] Inclusion of reported outcomes in their disaggregated item scores often performed better than total scores, however this is not useful for mapping as outcomes are usually only reported as total scores and are rarely reported with all individual factors. Non-linear functions showed a small improvement in model fit but add a greater level of complexity for the user of the mapping function. It was decided the improvement in model fit was not sufficient for this reduction in usability.

substantive differences in mapping. At the extreme ends of the range of response values mapping starts to break down. It is hence not recommended for populations solely at the upper and lower bounds of the response distribution. Figures B3 and B4 show the accuracy of these predictions graphically.

Table B2. Descriptive Statistics

	Mean	(SD)	Min	Max	Obs.	Floor	Ceiling	Skewness	Kurtosis
UKHLS									
Life Satisfaction	7.90	(1.88)	0.29	9.71	21,027	0.75%	34.79%	-1.27	4.85
SDQ	10.70	(5.77)	0	35	21,027	0.78%	0.01%	0.59	3.07
SDQ-I	4.67	(3.32)	0	20	21,027	6.51%	0.00%	0.86	3.52
Age	12.52	(1.69)	10	15	21,027				
Male (n, %)	10,492	(49.9%)							
Female (n, %)	10,535	(50.1%)							
ALS									
Life Satisfaction	6.50	(2.66)	0	10	337,657	4.14%	13.77%	-0.67	2.80
Happiness	6.62	(2.72)	0	10	337,657	4.33%	17.17%	-0.71	2.81
Age	12.57	(1.75)	11	16	337,657				
Male (n, %)	198,797	(45.52%)							
Female (n, %)	215,904	(49.44%)							

Notes: Data is from UK Household Longitudinal Study Youth Sampe (UKHLS) and Active Lives of Children and Young People Survey (ALS). Floor and Ceiling columns show the proportion of individuals reporting at the lowest and highest response categories, respectively.

Table B3. Model fits

SDQ

Full sample error statistics (robust OLS)

Without controls

	R-squared	RMSE	MAE	ME	AIC	BIC
Model 1	0.254991	1.622956	1.243217	-7.5E-10	80038.65	80054.56
Model 2	0.258633	1.619022	1.244787	5.94E-09	79937.61	79961.47
Model 3	0.258674	1.619016	1.243902	6.89E-09	79938.44	79970.25
Model 4	0.278076	1.597728	1.226773	-2.9E-10	79382.8	79422.57
Model 5	0.283994	1.591316	1.22374	-2.3E-10	79217.7	79289.28
Model 6	0.28499	1.590361	1.220658	-6.8E-10	79196.45	79299.85
Model 7	0.344339	1.519531	1.163287	3.56E-10	72908.28	73074.08

With controls

	R-squared	RMSE	MAE	ME	AIC	BIC
Model 1	0.277178	1.599024	1.217535	-9.7E-11	79424.92	79528.32
Model 2	0.281007	1.594822	1.216435	-4.3E-10	79315.24	79426.59
Model 3	0.28104	1.594823	1.21584	1.02E-09	79316.29	79435.59
Model 4	0.296857	1.57722	1.20258	4.1E-10	78850.52	78977.77
Model 5	0.302149	1.571423	1.199006	-4.1E-11	78699.66	78858.73
Model 6	0.302924	1.5707	1.19728	-9.4E-11	78684.32	78875.2
Model 7	0.357554	1.504557	1.145689	-3.1E-10	72526.4	72779.05

Tenfold cross-validation

Robust OLS

	MAE				RMSE			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Model 1	1.21833	0.01857	1.19009	1.24381	1.59909	0.04738	1.55420	1.68482
Model 2	1.21721	0.01566	1.19519	1.24636	1.59500	0.03619	1.54982	1.66663
Model 3	1.21679	0.02859	1.18087	1.25526	1.59508	0.03461	1.53480	1.67131
Model 4	1.14803	0.01770	1.11708	1.17357	1.50586	0.03525	1.46314	1.58733
Model 5	1.14765	0.01045	1.13048	1.16169	1.50444	0.03011	1.45247	1.56501
Model 6	1.14811	0.01603	1.12524	1.18138	1.50465	0.02889	1.44995	1.54637
Model 7	1.14806	0.02292	1.11254	1.18796	1.50483	0.02453	1.45281	1.53000

Tobit

	MAE				RMSE			
	Mean	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
Model 1	1.21839	0.02933	1.17486	1.26328	1.59967	0.02712	1.53948	1.63605
Model 2	1.21759	0.01555	1.19751	1.24558	1.59555	0.03080	1.54880	1.65518
Model 3	1.21714	0.01981	1.17824	1.24479	1.59536	0.03405	1.51568	1.65362
Model 4	1.14779	0.01679	1.11786	1.17425	1.50636	0.03880	1.46534	1.58157
Model 5	1.14804	0.02093	1.10751	1.17376	1.50499	0.03501	1.42724	1.54857
Model 6	1.14863	0.02379	1.11109	1.18960	1.50499	0.02735	1.46095	1.54334
Model 7	1.14835	0.01783	1.12035	1.17095	1.50517	0.01499	1.47810	1.52535

SDQ-I

Full sample error statistics (robust OLS)

Without controls

	R-squared	**RMSE**	**MAE**	**ME**	**AIC**	**BIC**
Model 1	0.23964	1.639591	1.265336	-2E-09	80467.51	80483.42
Model 2	0.243571	1.635386	1.260345	-3.4E-09	80360.51	80384.37
Model 3	0.243571	1.635425	1.260385	-1.7E-09	80362.51	80394.32
Model 4	0.240651	1.638539	1.265033	1.97E-09	80441.52	80465.38
Model 5	0.246066	1.632765	1.260236	-4.9E-09	80295.06	80334.83
Model 6	0.246794	1.632053	1.256825	5.49E-10	80278.72	80334.4
Model 7	0.304415	1.568306	1.208113	6.49E-10	76218.68	76305.83

With controls

	R-squared	**RMSE**	**MAE**	**ME**	**AIC**	**BIC**
Model 1	0.256269	1.621988	1.244149	7.75E-10	80024.55	80127.95
Model 2	0.260066	1.61788	1.240844	9.02E-11	79918.92	80030.27
Model 3	0.260066	1.617918	1.240879	9.78E-11	79920.91	80040.21
Model 4	0.257005	1.621223	1.244063	-1.1E-09	80005.72	80117.07
Model 5	0.261772	1.616091	1.240756	-2.4E-10	79874.39	80001.64
Model 6	0.262355	1.615529	1.239234	-1.1E-11	79861.77	80004.93
Model 7	0.317351	1.554075	1.190868	7.69E-10	75857.96	76032.26

Tenfold cross-validation

Robust OLS

	MAE				RMSE			
	Mean	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
Model 1	1.24515	0.02383	1.19012	1.27147	1.62229	0.04704	1.53978	1.70328
Model 2	1.24179	0.02239	1.21260	1.28132	1.61818	0.02437	1.58513	1.65960
Model 3	1.24197	0.02271	1.20706	1.27303	1.61815	0.04649	1.53117	1.68191
Model 4	1.24496	0.01861	1.21245	1.27289	1.62169	0.02096	1.58073	1.65250
Model 5	1.24171	0.02131	1.20900	1.27897	1.61664	0.03368	1.54300	1.64983
Model 6	1.24020	0.02360	1.18623	1.26440	1.61579	0.04794	1.53902	1.69142
Model 7	1.19238	0.01604	1.16648	1.21447	1.55485	0.03359	1.48844	1.60157

Tobit

	MAE				RMSE			
	Mean	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
Model 1	1.24513	0.01613	1.22010	1.26977	1.62226	0.03532	1.57017	1.67451
Model 2	1.24186	0.02267	1.20278	1.26960	1.61834	0.02167	1.57993	1.64359
Model 3	1.24191	0.01766	1.21125	1.27010	1.61854	0.02133	1.58753	1.64385
Model 4	1.24530	0.01761	1.22329	1.27134	1.62184	0.03011	1.56864	1.66281
Model 5	1.24218	0.02873	1.18259	1.27859	1.61688	0.04368	1.54941	1.69063
Model 6	1.24038	0.02970	1.18244	1.27727	1.61631	0.03275	1.56472	1.65141
Model 7	1.19245	0.02431	1.15961	1.23112	1.55536	0.03496	1.48772	1.60472

Happiness

Full sample error statistics (robust OLS)

Without controls

	R-squared	RMSE	MAE	ME	AIC	BIC
Model 1	0.414741	2.036212	1.562853	1.91E-09	1438245	1438267
Model 2	0.416504	2.033146	1.560966	-2.1E-09	1437228	1437261
Model 3	0.417607	2.031228	1.562316	2.37E-09	1436592	1436635

With controls

	R-squared	RMSE	MAE	ME	AIC	BIC
Model 1	0.422458	2.022777	1.550474	7.32E-11	1433786	1433925
Model 2	0.424447	2.019293	1.5494	1.18E-10	1432623	1432773
Model 3	0.425523	2.017409	1.549829	4.11E-10	1431994	1432155

Tenfold cross-validation

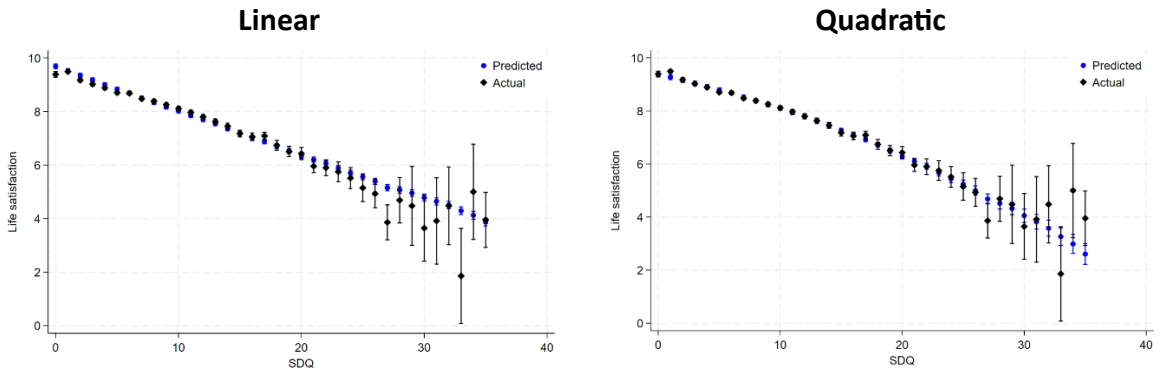
Robust OLS

	MAE				RMSE			
	Mean	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
Model 1	1.55054	0.00527	1.54275	1.55706	2.02282	0.00726	2.01275	2.03533
Model 2	1.54948	0.00918	1.53829	1.56390	2.01931	0.01238	2.00344	2.03769
Model 3	1.54990	0.00751	1.53504	1.55960	2.01748	0.01568	1.99339	2.04210

Tobit

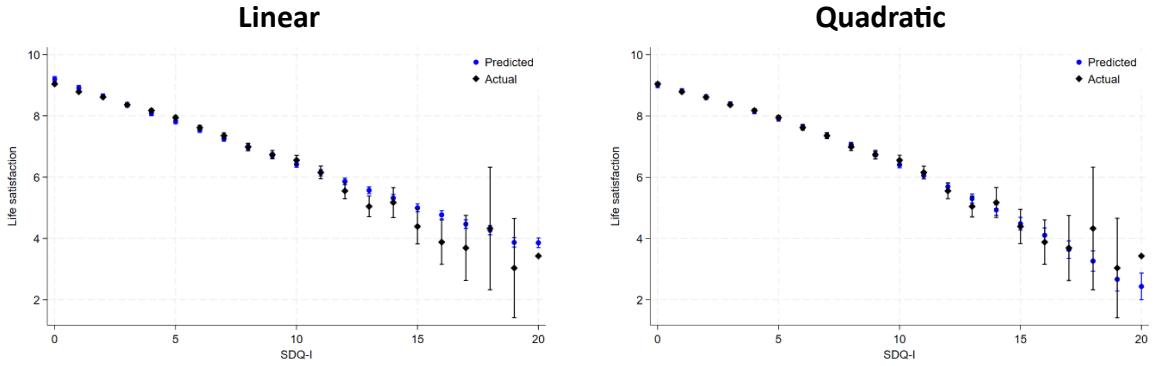
	MAE				RMSE			
	Mean	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
Model 1	1.52104	0.00879	1.51180	1.53462	2.04794	0.00544	2.03965	2.05873
Model 2	1.51981	0.00590	1.51246	1.53146	2.04535	0.01143	2.03358	2.06834
Model 3	1.51976	0.01184	1.50208	1.53612	2.04857	0.00742	2.03630	2.05873

Figure B3. SDQ Mapping Scatterplots



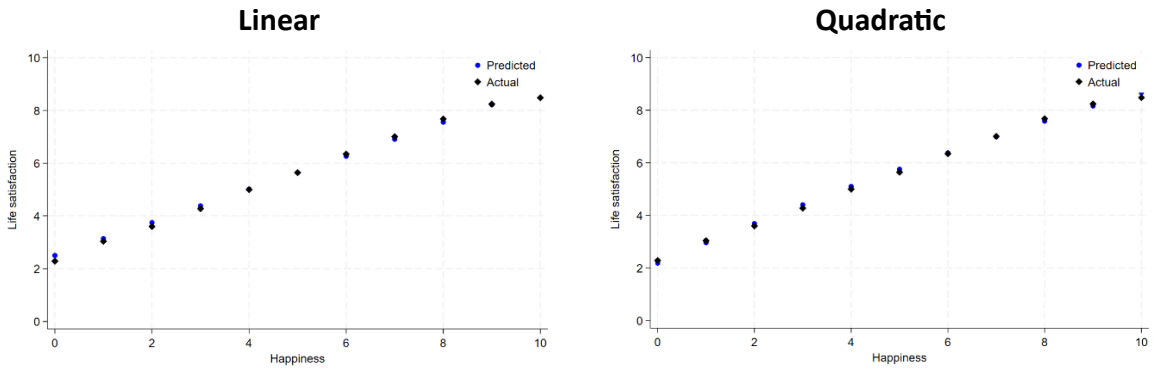
Notes: Actual values are the mean level of life satisfaction at each point on the SDQ scale. Whiskers show 95% confidence intervals. Predicted values are the mean predicted life satisfaction level for each level of SDQ based on the regression: $LifeSatisfaction_i = \alpha_1 + \theta_1 SDQ_i + X_i\beta + \varepsilon_i$ in the linear case and $LifeSatisfaction_i = \alpha_1 + \theta_1 SDQ_i + \theta_2 {SDQ_i}^2 + X_i\beta + \varepsilon_i$ in the quadratic case, where X_i is a vector of controls for age, gender, and year.

Figure B4. SDQ-I Mapping Scatterplots



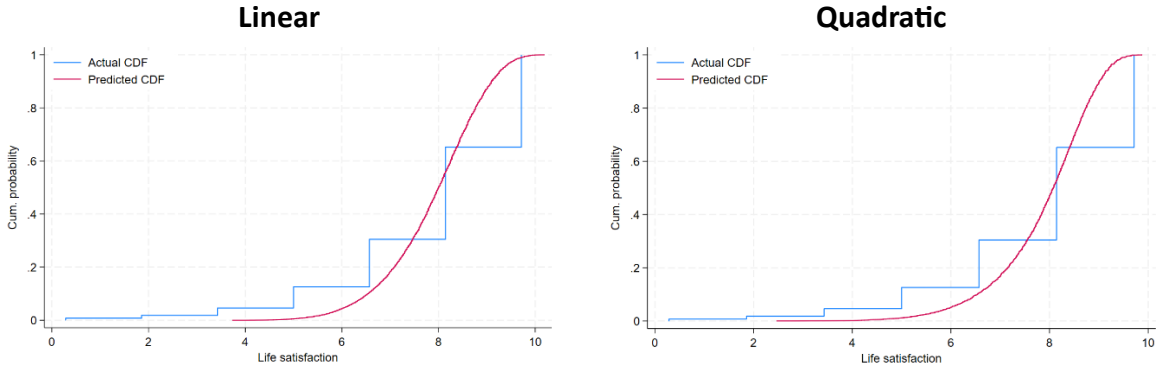
Notes: Actual values are the mean level of life satisfaction at each point on the SDQ-I scale. Whiskers show 95% confidence intervals. Predicted values are the mean predicted life satisfaction level for each level of SDQ-I based on the regression: $LifeSatisfaction_i = \alpha_1 + \theta_1 SDQ_I_i + X_i\beta + \varepsilon_i$ in the linear case and $LifeSatisfaction_i = \alpha_1 + \theta_1 SDQ_I_i + \theta_2 {SDQ_I_i}^2 + X_i\beta + \varepsilon_i$ in the quadratic case, where X_i is a vector of controls for age, gender, and year.

Figure B5. Happiness Mapping Scatterplots



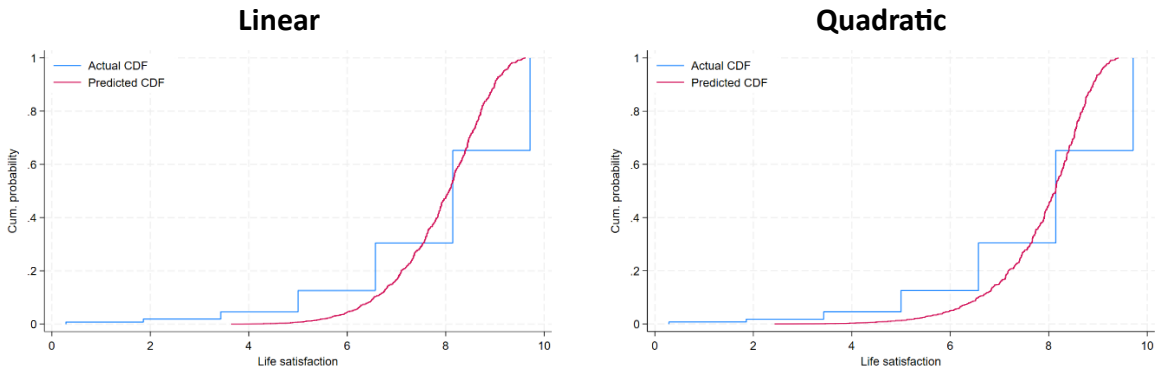
Notes: Actual values are the mean level of life satisfaction at each point on the Happiness scale. Whiskers show 95% confidence intervals. Predicted values are the mean predicted life satisfaction level for each level of Happiness based on the regression: $LifeSatisfaction_i = \alpha_1 + \theta_1 Happiness_i + X_i\beta + \varepsilon_i$ in the linear case and $LifeSatisfaction_i = \alpha_1 + \theta_1 Happiness_i + \theta_2 Happiness_i{}^2 + X_i\beta + \varepsilon_i$ in the quadratic case, where X_i is a vector of controls for age, gender, and year.

Figure B6. SDQ Mapping Cumulative Distribution Function



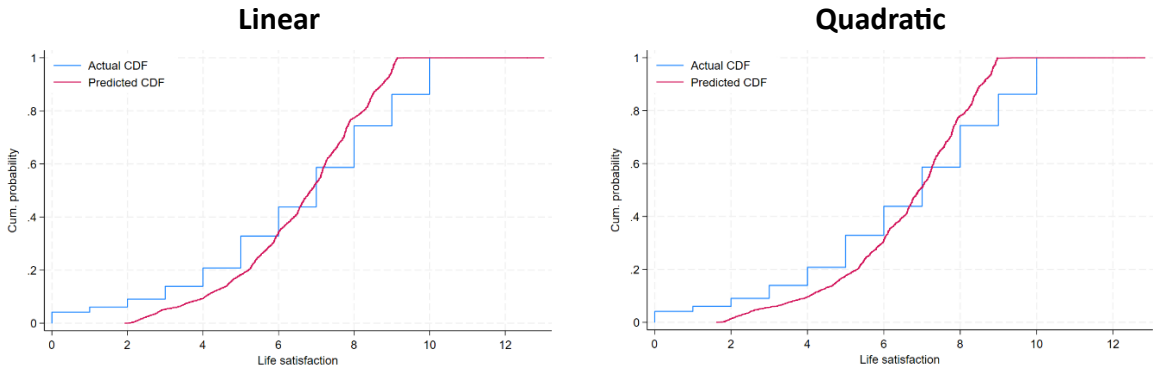
Notes: Actual CDF is the cumulative distribution function of life satisfaction responses. Predicted CDF is the predicted cumulative distribution function of life satisfaction responses based on the regression: $LifeSatisfaction_i = \alpha_1 + \theta_1 SDQ_i + X_i\beta + \varepsilon_i$ in the linear case and $LifeSatisfaction_i = \alpha_1 + \theta_1 SDQ_i + \theta_2 SDQ_i{}^2 + X_i\beta + \varepsilon_i$ in the quadratic case, where X_i is a vector of controls for age, gender, and year.

Figure B7. SDQ-I Mapping Cumulative Distribution Function



Notes: Actual CDF is the cumulative distribution function of life satisfaction responses. Predicted CDF is the predicted cumulative distribution function of life satisfaction responses based on the regression: $LifeSatisfaction_i = \alpha_1 + \theta_1 SDQ_I_i + X_i\beta + \varepsilon_i$ in the linear case and $LifeSatisfaction_i = \alpha_1 + \theta_1 SDQ_I_i + \theta_2 SDQ_I_i{}^2 + X_i\beta + \varepsilon_i$ in the quadratic case, where X_i is a vector of controls for age, gender, and year.

Figure B8. Happiness Mapping Cumulative Distribution Function



Notes: Actual CDF is the cumulative distribution function of life satisfaction responses. Predicted CDF is the predicted cumulative distribution function of life satisfaction responses based on the regression: $LifeSatisfaction_i = \alpha_1 + \theta_1 Happiness_i + X_i\beta + \varepsilon_i$ in the linear case and $LifeSatisfaction_i = \alpha_1 + \theta_1 Happiness_i + \theta_2 Happiness_i{}^2 + X_i\beta + \varepsilon_i$ in the quadratic case, where X_i is a vector of controls for age, gender, and year.