Construct validity, reliability and responsiveness of the 10-item Well-being instrument (WiX) for use in economic evaluation studies

EsCHER Working Paper No. 2023005
April, 2023

Judith Bom, PhD
Daphne Voormolen, PhD
Werner B.F. Brouwer, PhD
Esther de Bekker-Grob, PhD
Job van Exel, PhD

Erasmus University Rotterdam
Making Minds Matter

Title

Construct validity, reliability and responsiveness of the 10-item Well-being instrument (WiX) for use in economic evaluation studies

Authors

Judith Bom, PhD, Erasmus School of Health Policy & Management
Daphne Voormolen, PhD, Erasmus School of Health Policy & Management
Werner B.F. Brouwer, PhD, Erasmus School of Health Policy & Management
Esther de Bekker-Grob, PhD, Erasmus School of Health Policy & Management
Job van Exel, PhD, Erasmus School of Health Policy & Management

Corresponding author: Daphne Voormolen, voormolen@eshpm.eur.nl

Keywords

Well-being, measurement, outcomes, economic evaluation, validity, reliability, responsiveness, WiX

JEL classification

D04; D60; I31;

Cite as

Bom, J.A.M., Voormolen, D.C., Brouwer, W.B.F., de Bekker-Grob, E.W. and van Exel, J. (2023). Construct validity, reliability and responsiveness of the 10-item Well-being instrument (WiX) for use in economic evaluation studies. EsCHER Working Paper Series No. 2023005, Erasmus University Rotterdam. Available from: https://www.eur.nl/en/research/escher/research/working-papers

Erasmus Centre for Health Economics Rotterdam (EsCHER) is part of Erasmus University Rotterdam.
Want to know more about EsCHER? Visit www.eur.nl/escher
Want to contact EsCHER? E-mail escher@eur.nl

Interested in more EsCHER Working Papers? Download from www.eur.nl/escher/research/workingpapers

Abstract

Background Economic evaluations of interventions in health and social care require outcome measures that capture their full benefits, including those beyond health. The newly developed 10-item Well-being instrument (WiX) aims to capture the well-being of adults comprehensively yet concisely. The aim of this study is to assess the construct validity, test-retest reliability, and responsiveness of the WiX.

Methods Data was gathered via an online survey in a representative sample of the adult general population in the Netherlands (N=1,045). Construct validity was assessed by inspecting convergent, structural and discriminant validity, following the COSMIN methodology. Regression analyses of the WiX and its items on other validated measures of well-being were performed to assess the convergent validity of the instrument and the relevance of its items. Dimensionality of the WiX was assessed using exploratory factor analysis. To assess discriminant validity, several hypotheses in terms of well-being differences were assessed. Finally, a second survey was sent out two weeks after the initial survey (N=563; 53.9% response rate) to assess the test-retest reliability and responsiveness of the WiX.

Results The WiX showed to be correlated with alternative well-being measures as expected and able to sufficiently differentiate between relevant subgroups in the population. Moreover, the dimensionality analysis indicated that the WiX captures a broad array of elements relevant to well-being, including physical and mental health. The test-retest reliability was good, with an intraclass correlation coefficient of 0.82.

Conclusions The results regarding the construct validity, reliability and responsiveness of the WiX are favourable and indicate that this new instrument may be a promising alternative for existing measures of well-being for evaluating interventions in health and social care. Further validation of the measure in specific subgroups and in other countries is needed, and utility weights need to be determined before the WiX can be used in economic evaluations.

Acknowledgement

Erasmus Initiative Smarter Choices for Better Health, Erasmus University Rotterdam (https://www.eur.nl/en/research/research-initiatives/erasmus-initiatives/smarter-choices-better-health)

Introduction

Increasing health and social care expenditures worldwide emphasize the need for optimal allocation of scarce resources in these areas. Economic evaluations can aid such decisions by identifying, measuring, valuing, and comparing the costs and benefits of the interventions that are considered (Drummond et al., 2005). The benefits of interventions in the health domain are commonly captured in terms of Quality-Adjusted Life-Years (QALYs), which comprise both length of life and health-related quality of life (HRQoL). The EuroQol-5D (EQ-5D) instrument (EuroQol Group, 1990), the Short-Form Health Survey (SF-36) (Ware & Sherbourne, 1992) or the Health Utilities Index (HUI) (Furlong et al., 2001) are often recommended and used to measure HRQoL in the context of economic evaluations in healthcare (Kennedy-Martin et al., 2020; Wisløff et al., 2014). Measures that focus on HRQoL may, however, fall short in capturing the full benefits of an intervention, when the intervention does not (only) improve physical or mental health but (also) impacts broader aspects of quality of life. Using incomplete information about the benefits of an intervention for economic evaluations may eventually lead to suboptimal allocation decisions. Hence, it has been argued that especially for interventions in long-term care, social care, and palliative care, broader outcome measures are required (e.g., (Coast, 2014; Johnson et al., 2016; Makai et al., 2014)). The need for such broader measures may be even more evident in the context of prevention policies or intersectoral interventions, aiming to improve health as well as other elements of well-being.

This need for broader outcome measures that can be used in economic evaluations of interventions in health and social care has resulted in the introduction of several well-being instruments over the past years. For instance, the ICEpop CAPability (ICECAP) instruments, which aim to capture capability well-being in specific target groups (i.e., ICECAP-A for adults and ICECAP-O for older people) (Al-Janabi et al., 2012; Coast et al., 2008; Grewal et al., 2006), the Adult Social Care Outcomes Toolkit (ASCOT), specifically aimed at social care users and their caregivers (Netten et al., 2012), and the Well-being of Older People (WOOP) instrument, aimed at capturing experienced well-being in older people (Mariska Q. N. Hackert et al., 2021). In addition, the EuroQol Group has recently presented the EuroQol Health and Well-being (EQ-HWB) instrument, which aims to measure health-related well-being in the adult population (J. Brazier et al., 2022). All these measures come with its own strengths and weaknesses: their theoretical embedding is difficult to discern, they do not measure well-being comprehensively, or are confined to a specific subgroup of the population. For example, the WOOP and the ICECAP-O focus on older people and the ASCOT and EQ-HWB on care users and carers. In addition, the ICECAP instruments may not fully capture health effects of interventions, also not indirectly (M. Q. N. Hackert et al., 2017; Keeley et al., 2016). Therefore, these instruments may miss elements of value to the well-being of members of the general population that ideally would be included in evaluations of interventions in health and social care.

The 10-item Well-being instrument (WiX) was developed to improve welfare economic evaluations in the field of health and social care, while also allowing comparisons with interventions in other sectors, or across sectors. The WiX was grounded in existing well-being theories and empirical evidence of what the general adult population in the Netherlands considers to be important constituents for their well-being (Van der Deijl et al., 2023). This new instrument, therefore, aims to measure the functionings of members of the adult population on important domains of well-being. A content validation study has shown that the WiX is relevant, comprehensive and comprehensible (Voormolen et al., 2023). The aim of this study is to assess the construct validity, reliability, and responsiveness

of the WiX following the COSMIN methodology (Mokkink et al., 2018) in a representative sample of the adult population of the Netherlands.

Methods

Sampling strategy

Data was gathered in two stages between 7-12 October 2021 and 23-28 October 2021 via two separate online surveys. For the first survey (main sample), a sampling agency recruited 1,045 respondents, quota-sampled to be representative for the adult population of the Netherlands based on age, sex, and level of education. In this survey, respondents were asked about their background characteristics and their well-being and health using a variety of measures (see below). Two weeks after completing the first survey, respondents were invited to participate in a follow-up survey (retest sample). This second survey was used to test the reliability (in terms of test-retest) and responsiveness of the WiX. The time interval of two weeks was chosen following COSMIN recommendations (Mokkink et al., 2018) as being long enough to prevent recall bias while short enough to ensure limited individual changes in individual well-being. Over half of the main sample replied to this second survey (N=563; 53.9%). Participation to both surveys was voluntary and respondents received a small financial compensation. As all questions were mandatory, there were no missing answers.

Measures

The WiX aims to capture the overall well-being of adult members of the general population by asking them to assess how satisfied they are today on ten domains that were identified as most relevant for well-being in this group (Voormolen et al., 2023). Each item offers five response levels: “I’m very satisfied”; “I’m satisfied”; “I’m not satisfied but also not dissatisfied”; “I’m dissatisfied”; and “I’m very dissatisfied”. By attaching scores to these levels ranging from 1 “very dissatisfied” to 5 “very satisfied” and aggregating these scores over all items, a total score for the WiX can be computed that ranges from 10 (lowest level of well-being) to 50 (highest level of well-being). The English version of the instrument can be found in Appendix A. Respondents were asked to complete the WiX in both surveys. In the follow-up survey they were also asked to report whether significant events affecting their well-being -positively and/or negatively- had occurred in the time between answering to the two surveys and, if so, to describe these event(s).

In addition to the WiX, well-being was measured using the Satisfaction with Life Scale (SWLS) and the Cantril Ladder, which are measures with a similar focus as the WiX on experienced well-being. The SWLS aims to capture individuals’ global judgement of their life, based on five items with seven response levels each, leading to an aggregate score ranging from 5 (lowest satisfaction with life) to 35 (highest satisfaction with life) (Diener et al., 1985). The Cantril Ladder (Cantril, 1965) is a vertical visual analogue scale (VAS) that represents a ladder of life ranging from 0 (worst possible life) to 10 (best possible life). Individuals were asked to indicate where on this ladder they felt they were standing at the moment.

The health of respondents was measured using the five-level EQ-5D instrument (EQ-5D-5L) and the EuroQol Visual Analogue Scale (EQ-VAS). The EQ-5D-5L (Herdman et al., 2011) aims to measure and value HRQoL on five dimensions. For each of these dimensions, individuals indicate the severity of their experienced issues using one of five response categories. Using utility weights for the Netherlands, utility scores ranging between -0.446 and 1 can be calculated (Versteegh et al., 2016), with 1 representing perfect health, 0 representing dead, and utility scores lower than 0 representing health states considered worse than being dead. We also included the cognition bolt-on for the EQ-5D to the survey, which aims to capture issues with concentration and memory (Krabbe et al., 1999). Furthermore, respondents were asked to rate their overall health today on a vertical scale ranging from 0 (worst imaginable health) to 100 (best imaginable health) (Herdman et al., 2011).

Background characteristics

Furthermore, the following background characteristics were included in the first survey: age; sex; household composition; educational attainment, classified as low (no, primary, pre-vocational education), middle (secondary or middle vocational education), or high education (higher vocational or academic education); work status; and household income, using a closed question with income intervals and a question asking how well the household can make ends meet financially, with four response categories (with great difficulty, with difficulty, fairly easily and easily).

Analytical strategy

This study aimed to assess the construct validity, reliability, and responsiveness of the WiX following the definitions and guidelines for assessment of instrument development presented in the COSMIN framework (Mokkink et al., 2010). In the main analyses all responses were incorporated. To assess the robustness of the results to speeders, the analyses were also run excluding the 5% fastest respondents. This did not significantly affect the findings from the analyses or their interpretation.

Descriptive statistics were computed for all measures. In addition, the correlations between WiX items were inspected to assess their relevance. As the WiX was designed as a multi-dimensional measure of well-being, the items were expected to be positively correlated but not highly, as they supposedly represent distinct domains of well-being. Very high correlation between items could indicate redundancy. Following guidelines of Hopkins (2002) strength of correlation was evaluated as follows: < 0.10, trivial; 0.10-0.29, small; 0.30-0.49, moderate; 0.50-0.69, high; 0.70-0.89, very high; ≥ 0.90, (nearly) perfect (Hopkins, 2002).

The construct validity of a measure encompasses three types of validity: convergent validity, structural validity and discriminative/known-groups validity. Convergent validity concerns the degree to which the WiX is related to instruments that aim to measure the same concept. This was assessed by inspecting the correlation of (the items of) the WiX with two alternative validated measures of experienced well-being (Cantril Ladder & SWLS). For ease of interpretation, the scores on these measures were assumed to be continuous and, hence, OLS-regression estimates were used. WiX (item) scores were expected to be positively and highly correlated with scores on these well-being measures. Lastly, we assessed the correlation between the WiX and two health measures (EQ-VAS and

EQ-5D-5L). Here we expected high correlations of the scores on the two health-related WiX items -'Mental health' and 'Physical health'- with scores on these measures, and lower but still positive correlations for scores on the other WiX items.

Structural validity concerns whether the scores of the WiX adequately reflect the dimensionality of the construct. This was analysed using exploratory factor analysis (EFA). Specifically, we assessed the overlap in factors between the WiX and the EQ-5D-5L. We expected the items of the WiX (covering more than health alone) to correspond with a larger set of factors than the items of the EQ-5D-5L. Bartlett's test of sphericity (Bartlett, 1951) was used to ensure that the correlation matrix was not random, and the Kaiser-Meyer Olkin (KMO) Measure of Sampling Adequacy (Kaiser, 1974) was applied to inspect whether the data was suited for factor analysis. As the items of both measures concern ordinal variables, we first calculated a polychoric correlation matrix to use for the principal factor analysis. The appropriate number of factors was selected based on the Kaiser criterium, the scree plot and the interpretability of the models. To allow the factors to be correlated to each other, oblique rotation was applied. Promax was used for the main analysis, oblimin rotation was used as a robustness check. Furthermore, in our main model we only considered the original EQ-5D-5L items. As a robustness check, the analysis was repeated including the cognition bolt-on.

Discriminative or known-groups validity concerns whether the instrument can discriminate between relevant subgroups in the sample. This was investigated by inspecting whether scores on the WiX (items) differed between selected subgroups as expected using t-tests and One-Way Analyses of Variance (ANOVA). Based on previous research, we expected respondents of higher age, with higher level of education, who are employed, with higher income, able to make ends meet (fairly) easily, and those in an intimate relationship to report higher levels of well-being (Dolan et al., 2008). Additionally, for those individuals with financial difficulties, we expected lower scores for the WiX item 'Financial situation'. Likewise, for those individuals in an intimate relationship we expected higher scores on the item 'Relations'. Moreover, previous literature showed strong associations between health and well-being (Dolan et al., 2008). Hence, we expected individuals in poor health based on their EQ-5D-5L scores to have lower total WiX scores as well as lower scores on the WiX items 'Mental health' and 'Physical health'. In addition, in previous research, cognition was found to be an important element of well-being (e.g., (Davis et al., 2015)). Consequently, we expected lower WiX scores among those who reported issues with cognition.

The reliability and responsiveness of the WiX were assessed using the data from the retest sample. To evaluate the test-retest reliability of the WiX, percentages of complete agreement and quadratic weighted kappa statistics were calculated for all items and the overall measure. Complete agreement represents the share of respondents that reported the exact same score at both time points (t0 and t1). The kappa statistic measures agreement and is scaled to 0 when the amount of agreement observed could have been expected due to chance and 1 when perfect agreement is observed, while accounting for variation in inconsistent responses due to the ordinal nature of our data. As a robustness check, we calculated kappa statistics using linear weights. Following Landis & Koch (1977), the following interpretation of the scores were applied: 0.21-0.40, fair; 0.41-0.60 moderate; 0.61-0.80 substantial; 0.81-1.00, (nearly) perfect (Landis & Koch, 1977). The consistency in responses on total WiX scores was evaluated by calculating intra-class correlation (ICC) using a two-way mixed effects model (Koo & Li, 2016). Following their suggestions, reliability scores were interpreted as follows: 0.5-0.75, moderate; 0.75-0.90, good; >0.90, excellent.

Using the retest sample, we also evaluated the responsiveness of the WiX based on the well-being scores of respondents who reported to have experienced an event that significantly affected their well-being in the two-week time-interval between the two surveys (N=247). Responsiveness was assessed by investigating whether there was a significant difference in the change in well-being scores between those respondents who faced a shock (i.e., a positive or negative event as interpreted by the respondent themselves) and those who did not.

Ethics

The study protocol was approved by the Research Ethics Review Committee of the Erasmus School of Health Policy & Management (case number 21-001). Participation in the study was voluntary and could be terminated at any point. All respondents provided written informed consent for participation in this study and use of their responses for scientific research purposes.

Results

Sample characteristics

Descriptive characteristics of our main sample and the retest sample can be found in Table 1. The main sample was representative for the adult general population of the Netherlands in terms of age, sex, and education. Respondents in the retest sample were more often female and were older.

Table 1: Descriptive statistics main sample and retest sample

Variables	Main sample (N=1,045)	Retest sample (N=563)	Difference
Sex			
Male	50.1%	43.9%	*
Female	49.8%	56.1%	
Other/Prefer not to tell	0.1%	-	
Age			
18-24	11.1%	6.9%	***
25-34	17.9%	11.0%	
35-44	22.0%	24.2%	
45-54	21.8%	27.7%	
55-70	27.2%	30.2%	
Education			
Low	33.0%	31.3%	
Middle	42.1%	41.9%	
High	24.9%	26.8%	

Note: * p<0.05, ** p<0.01, *** p<0.001. Education is categorized into Low (no, primary, pre-vocational education); Middle (secondary or middle vocational education); High (higher vocational or academic education).

Figure 1 shows the distribution of scores on the items of the WiX (Table B1 in Appendix B presents the underlying scores). The majority of respondents from the main sample indicated to be satisfied or very satisfied on all items of the instrument, with the lowest proportion of individuals indicating to be (very) satisfied with their physical health (53.9%) and financial situation (56.1%). Few individuals were (very) dissatisfied with their safety (5.5%) or living environment (7.6%). The resulting mean total WiX score in the main sample was 37.5 (SD 6.84); Figure 2 shows the distribution of these scores. Although quite a large proportion of respondents reported to be (very) satisfied on most of the items, the ceiling effect appears to be moderate when looking at the total WiX scores: 5.2% of the respondents reported to be very satisfied on all items of the WiX and had a total WiX score of 50. Only one respondent indicated to be very dissatisfied on all items of the WiX.

Table 2 depicts the correlations between scores on the items of the WiX. Most items were moderately correlated to each other, at most, supporting the relevance of the separate items. High correlations were found for 'Physical health' with 'Mental health' and 'Self-worth'; for 'Safety' with 'Living environment'; for 'Activities' with 'Relaxation and leisure time'; for 'Independence' with 'Relaxation and leisure time' and 'Self-worth'; and for 'Self-worth' with 'Mental health', 'Relaxation and leisure time', 'Activities', and 'Independence'.

Figure 1: Distribution of responses on the 10 items of the WiX (N=1,045)

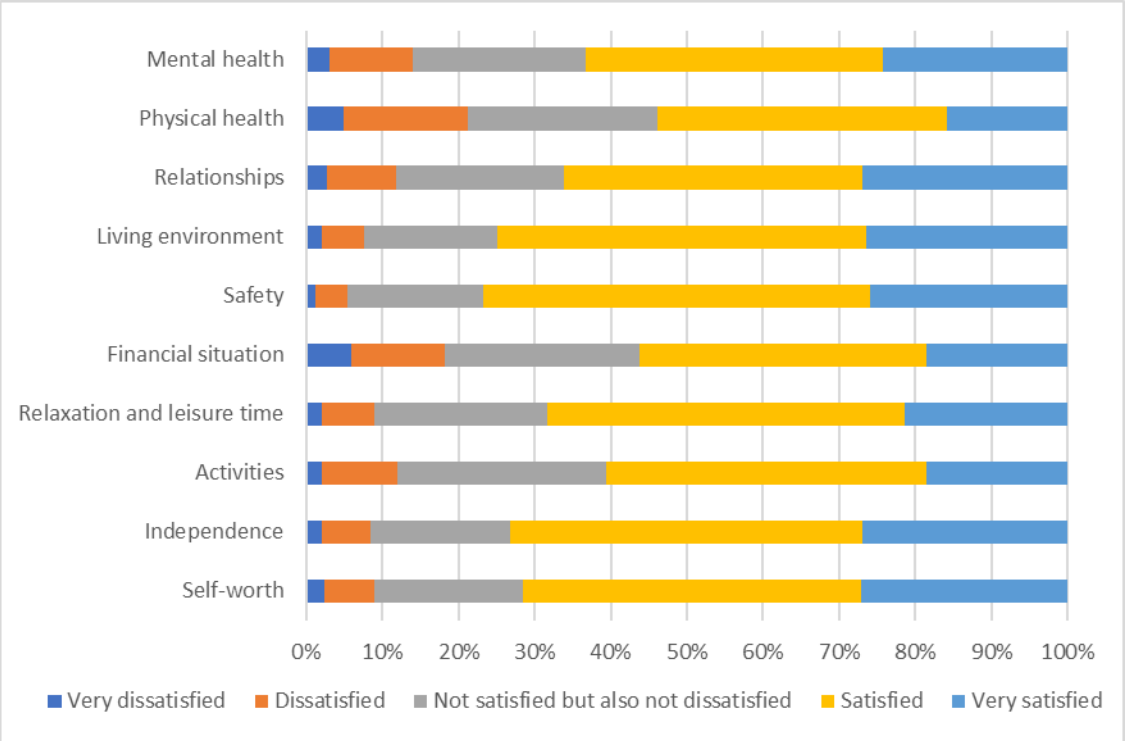


Figure 2: Distribution of total WiX scores in the main sample (N=1,045)

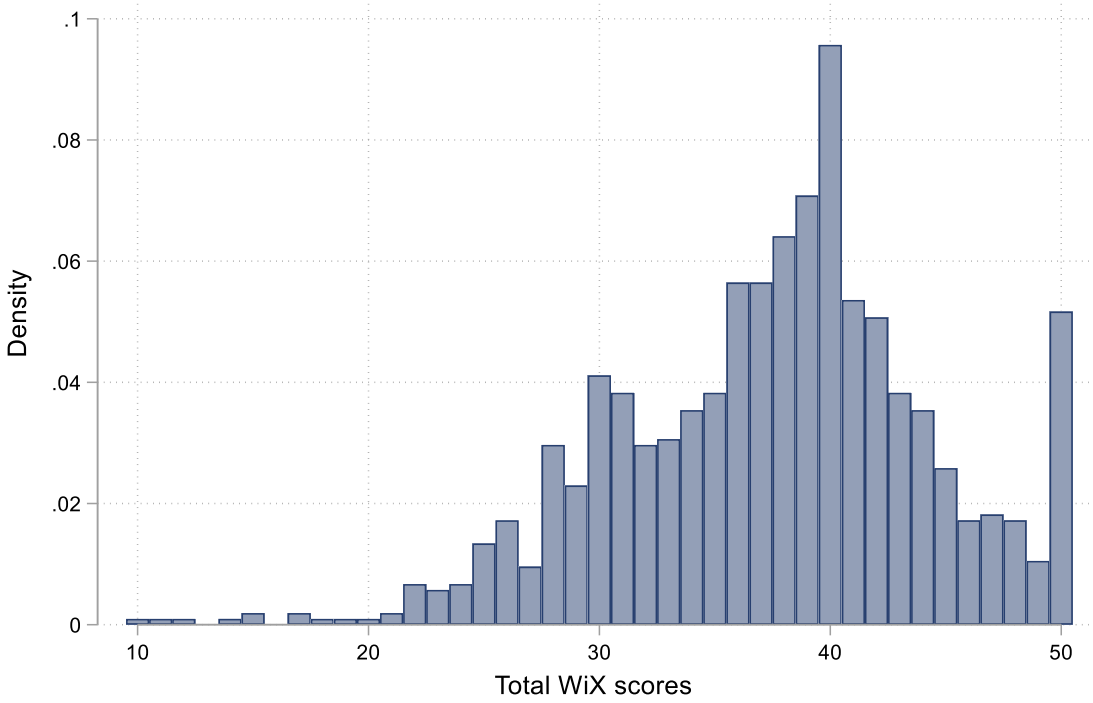


Table 2: Spearman correlations of items of the WiX (N=1,045)

	Mental health	Physical health	Relationships	Living environment	Safety	Financial situation	Relaxation and leisure time	Activities	Independence	Self-worth
Mental health	1									
Physical health	0.539***	1								
Relationships	0.431***	0.332***	1							
Living environment	0.338***	0.292***	0.445***	1						
Safety	0.384***	0.337***	0.387***	0.529***	1					
Financial situation	0.439***	0.433***	0.374***	0.391***	0.378***	1				
Relaxation and leisure time	0.426***	0.385***	0.386***	0.467***	0.421***	0.458***	1			
Activities	0.474***	0.454***	0.436***	0.412***	0.425***	0.474***	0.544***	1		
Independence	0.405***	0.375***	0.386***	0.430***	0.436***	0.412***	0.524***	0.475***	1	
Self-worth	0.572***	0.420***	0.492***	0.418***	0.444***	0.480***	0.502***	0.524***	0.536***	1

* p<0.05, ** p<0.01, *** p<0.001

Convergent validity

The correlations of (the items of) the WiX with the Cantril Ladder and SWLS were moderate to high (Table 3). Regression analysis of the items of the WiX on both these subjective well-being measures, while controlling for individual characteristics (age, sex, education), indicated that all items of the WiX were positively associated with scores on both measures (Tables B2 and B3 in Appendix B). When adding all items in one model, nearly all items were still positively associated with scores on both measures but not all associations were statistically significant.

Additionally, correlations of scores on (the items of) the WiX with scores on two health measures (EQ-5D-5L and EQ-VAS) showed high correlations for the health-related items of the WiX, especially for 'Physical health' (Table 3). Correlations with other items of the WiX, such as 'Living environment' and 'Safety', were small. Table B4 (in Appendix B) shows correlations between the items of the WiX and the items of the EQ-5D-5L; correlations were especially strong for WiX item 'Physical health' with EQ-5D-5L item 'Pain and discomfort', and for WiX item 'Mental health' with EQ-5D-5L item 'Anxiety & depression'.

Table 3: Spearman correlations of items of the WiX with measures of well-being and health (N=1,045)

	Cantril Ladder	**SWLS**	**EQ-VAS**	**EQ-5D-5L**
Mental health	0.491***	0.399***	0.486***	0.477***
Physical health	0.480***	0.307***	0.623***	0.544***
Relationships	0.388***	0.374***	0.302***	0.291***
Living environment	0.324***	0.300***	0.261***	0.212***
Safety	0.333***	0.266***	0.277***	0.267***
Financial situation	0.521***	0.388***	0.408***	0.320***
Relaxation and leisure time	0.396***	0.358***	0.338***	0.303***
Activities	0.451***	0.357***	0.401***	0.382***
Independence	0.355***	0.341***	0.347***	0.336***
Self-worth	0.445***	0.406***	0.423***	0.396***
WiX total	0.613***	0.548***	0.561***	0.532***

Notes: * p<0.05, ** p<0.01, *** p<0.001. EQ-5D-5L, five-levels EuroQol five-dimensional questionnaire; SWLS, Satisfaction with Life Scale; EQ-VAS, EuroQol Visual Analogue Scale

Structural validity

Using EFA to assess the overlap in factors between the WiX and the EQ-5D-5L, we identified three factors (Table 4). The first factor contained only items of the WiX, while the other two were a combination of both WiX and EQ-5D items. Factor 2 seems to capture items related to mental health and factor 3 items related to physical health.

Table B5 and B6 (in Appendix B) present the results of two alternative models (i.e., using oblimin rotation instead of promax & adding the EQ-5D bolt-on cognition to the model), which show slightly different results but also clearly indicate that the WiX seems to capture dimensions of well-being beyond health.

Discriminative/known-groups validity

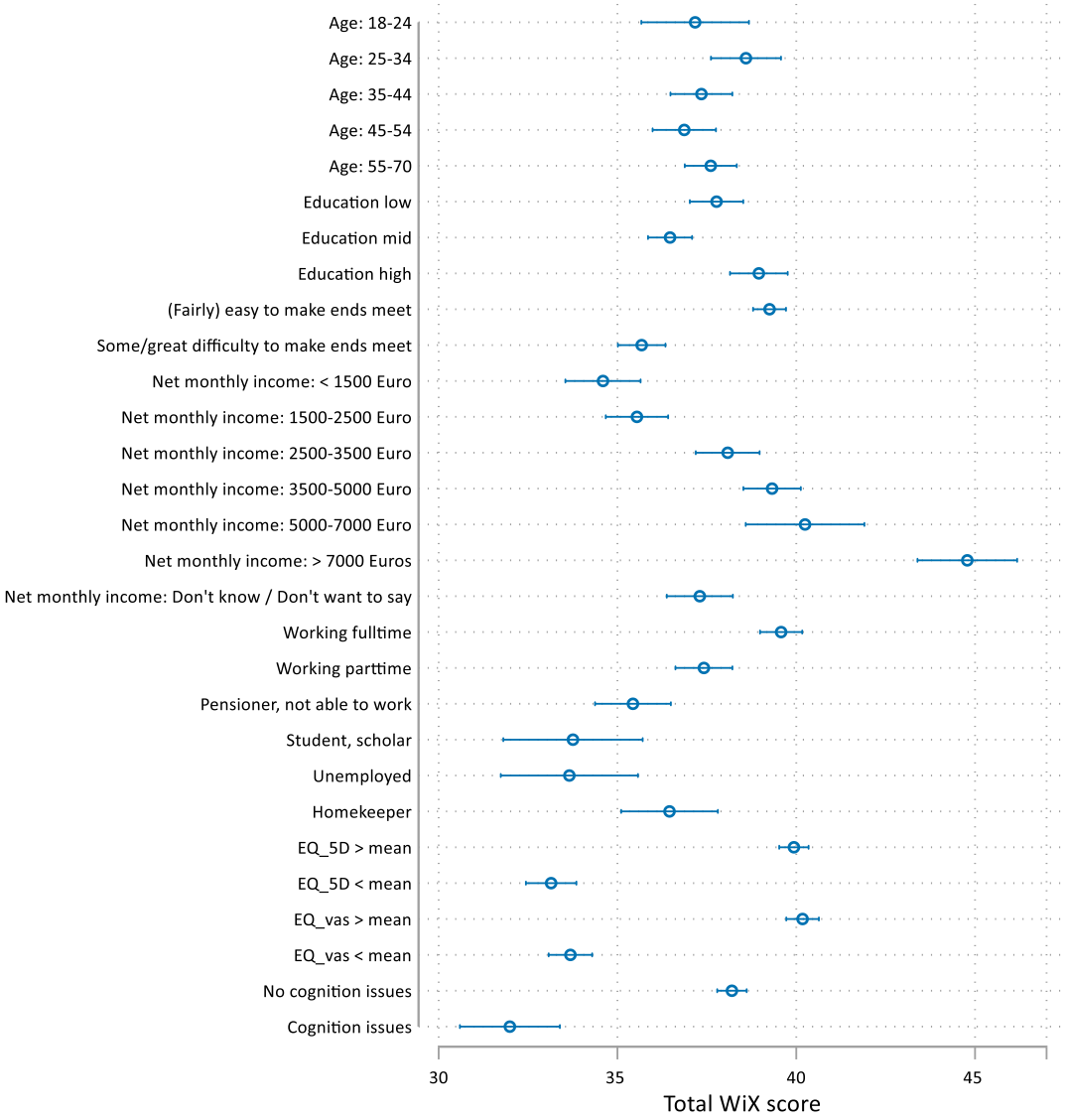
Figure 3 graphically depicts the mean WiX scores for subgroups in the sample. Except for age, we found WiX scores to differ as expected between subgroups. Significantly higher WiX (item) scores were observed among respondents with higher education, making ends meet (fairly) easily, living together with a partner, being employed and being in better health (Tables B7 and B8 in Appendix B present the underlying estimates).

Table 4: Factor loadings for the items of the WiX and the EQ-5D-5L (N=1,045)

Variable	Factor1	Factor2	Factor3	Uniqueness
Mental health		0.751		0.292
Physical health			0.384	0.305
Relationships	0.516			0.543
Living environment	0.801			0.444
Safety	0.715			0.476
Financial situation	0.579			0.497
Relaxation and leisure time	0.688			0.398
Activities	0.576			0.402
Independence	0.658			0.432
Self-worth	0.473	0.458		0.328
Mobility			0.924	0.249
Self-care			0.823	0.275
Usual activities			0.806	0.187
Pain and discomfort			0.760	0.277
Anxiety & depression		0.641		0.369

Note: Promax rotation, factor loadings below 0.3 are dropped from the table to allow easy interpretation of results

Figure 3: Overview of average total WiX scores by subgroups



Note: Bars indicate 95% confidence intervals.

Reliability

In Table 5 we show the test-retest reliability results when limiting our sample to respondents who indicated not to have experienced an event that affected their well-being significantly in the two-week time-interval between participating in the two surveys (N=316; 56.1%) (Table B9 in Appendix B shows the results for the full retest sample (N=563), which were very similar; the results of the robustness check where linear weights were used were also fairly similar (Table B10)). The correlations between WiX (item) scores in the main and retest samples were mostly high to very high. The percentages complete agreement, representing the proportion of respondents with the exact same score on the WiX (items) at t0 and t1, ranged from 57.0% to 69.0% for the items of the WiX and was 15.5% for the total WiX score. The weighted Kappa scores for the WiX items ranged from 0.48 to 0.80, which indicates moderate to substantial agreement (Landis & Koch, 1977), while the ICC score of 0.82 for the total WiX scores can be interpreted as good reliability (Koo & Li, 2016).

Table 5: Test-retest reliability estimates for stable retest sample (N=316)

	Correlations main and retest sample	Complete agreement	Weighted Kappa		
			mean	lower bound	upper bound
Mental health	0.686***	61.1%	0.70	0.63	0.77
Physical health	0.704***	60.8%	0.70	0.62	0.77
Relationships	0.666***	63.3%	0.69	0.61	0.76
Living environment	0.630***	68.0%	0.59	0.46	0.68
Safety	0.519***	65.2%	0.48	0.36	0.59
Financial situation	0.775***	69.9%	0.80	0.75	0.84
Relaxation and leisure time	0.568***	63.0%	0.57	0.47	0.66
Activities	0.518***	57.0%	0.53	0.43	0.62
Independence	0.735***	69.0%	0.72	0.65	0.79
Self-worth	0.686***	64.6%	0.73	0.65	0.79
WiX total	0.809***	15.5%	0.82[a]	0.79	0.86

Note: [a] For the total (continuous) WiX score two-way intra-class correlation is reported instead of weighted kappa.

Responsiveness

In the retest sample, 247 respondents (43.9%) reported an event that affected their well-being significantly in the two weeks between participating in the first and the second survey. Among this sample, 115 respondents reported to have experienced an event that positively affected their well-being, 74 respondents an event that negatively affected their well-being, and 58 respondents reported both. Respondents who experienced a positive event most often mentioned to have been on holidays or to have experienced something pleasant such as meeting a friend/family member or going out for dinner, while those who experienced a negative event most respondents mentioned something related to their mental or physical health. Compared to those not experiencing an event within the two-week interval, experiencing a positive event showed to be associated with a statistically significant increase in the total WiX score of 1.24. The change in WiX score did not significantly differ between those experiencing a negative event and those not experiencing any impactful event.

Discussion

To be able to assess outcomes broader than health in evaluations studies there is a need for outcome measures capturing overall well-being comprehensively. This study assessed the construct validity, reliability, and responsiveness of the WiX, a newly developed instrument to comprehensively measure the functionings of adults on ten important domains of well-being. Analyses were carried out in a representative sample of 1,045 members of the adult general population of the Netherlands, of whom 563 (53.9%) also participated in a second retest survey two weeks after completion of the first. The analyses demonstrated that the items of the WiX appear to be relevant, as the ten items of the WiX were positively but not too strongly correlated with each other, and that the construct validity and test-retest validity of the instrument are sufficient.

Correlations between the WiX (items) and other well-being measures with a similar focus (SWLS and the Cantril Ladder) were positive and high. Regression analyses showed that higher scores on the items of the WiX were positively associated with scores on both these subjective well-being measures. However, when including all the WiX items in these models simultaneously, some associations were not statistically significant. This could relate to limited variation in well-being scores in our sample, with few respondents reporting (very) poor well-being on most items. For example, about 1% of the sample indicated to be highly dissatisfied with their safety. Furthermore, the variation in associations of WiX items with the two alternative subjective well-being measures may partly reflect differences in the concepts these measures capture. For example, the WiX items 'Mental health', 'Physical health', 'Self-worth', and 'Financial situation' were strongly associated with both subjective well-being measures, but the SWLS also correlated with the WiX items 'Relationships' and 'Living environment' whereas the Cantril Ladder also correlated with 'Activities'. To further assess the relevance of specific WiX items in measuring well-being, future research should focus on studying subgroups of individuals who experience low levels of well-being and are expected to show more variation in scores on the WiX items. For instance, studying whether households experiencing financial hardship indeed indicate to be dissatisfied with their financial situation more often and investigating how this relates to their overall well-being would provide more insight in the relevance of this domain, also in the lower scoring levels.

Using EFA, the items of the WiX showed to relate with items of the EQ-5D-5L in the expected way, which indicates structural validity. The dimensionality analysis identified three factors: two factors capturing the health-related items of the WiX and domains of the EQ-5D, representing mental and physical health dimensions of well-being, and a third factor comprising the non-health items of the WiX. The latter can be interpreted as broader elements considered important for well-being that are not covered by this HRQoL measure (but possibly also not by its extended version, the EQ-HWB; (J. Brazier et al., 2022)). These results align with earlier studies comparing the EQ-5D with well-being measures that also indicated that well-being measures capture elements beyond health, not captured by the EQ-5D (Mariska Q. N. Hackert et al., 2021; Keeley et al., 2016). The results of the dimensionality analysis also emphasize that, as intended, the WiX is a comprehensive measure capturing both (mental and physical) health as well as broader elements of well-being.

When comparing WiX scores between subgroups, the observed differences were largely in line with our expectations based on previous research (Dolan et al., 2008). Individuals with a higher education, who are employed, make ends meet (fairly) easily, live with a partner, and are in better health, reported higher well-being scores. We did not observe a difference in well-being between individuals in different age groups, while this was expected based on previous research reporting a U-shaped relationship between age and well-being (Dolan et al., 2008). However, it has been argued that the presence of such a relationship might be driven by other elements associated with old age such as a decline in health (Ulloa et al., 2013). Additionally, there is some evidence that the hypothesized increase in well-being might mainly occur in individuals aged 70 and above (Hansen & Slagsvold, 2012), who were not included in our sample. The timing of our study (i.e., during the COVID pandemic) may also have influenced this result.

A retest-sample comprising 563 individuals showed sufficient test-retest reliability of the WiX. The overall weighted kappa scores of the items showed moderate to substantial agreement and the reliability of the total score showed to be good. The weighted kappa scores were lowest for the items 'Safety', 'Activities', 'Living

environment' and 'Relaxation and leisure time', and highest for the items 'Mental health', 'Physical health' and 'Financial situation'. These results might reflect the less stable nature of items like 'Relaxation and leisure time' as compared to health and finances when measured over a two-week interval. Earlier research indicated that physical constructs often show to be more reliable than social ones (J. E. Brazier et al., 1992).

Evidence regarding the responsiveness of the WiX is still limited. Only a small proportion of the retest sample mentioned to have experienced an event (positive or negative) significantly affecting their well-being in the two weeks in-between participating in the two surveys. Among those respondents, we found a significant increase in the WiX score for those experiencing positive events, but not a significant decrease for negative ones. These findings may relate to the type of events reported by respondents, which mostly concerned relatively minor events (in terms of a well-being 'shock') like a visit to the theatre or having a cold. Future (intervention) studies in larger samples experiencing potentially more impactful events are needed to assess the responsiveness of the WiX more adequately.

While the analyses presented here indicated sufficient validity of the WiX in a sample of the adult general population, the validity of the WiX among people experiencing poor well-being deserves further attention. Moreover, the timing of our study (i.e., during the COVID pandemic) as well as the use of an online survey should be considered in interpreting our findings. It seems relevant to explore the psychometric properties of the WiX (i.e., feasibility, construct validity, reliability and responsiveness) further in subgroups of the population that are expected to report low levels of well-being and using pen and paper or an interview-based administration of the instrument.

Moreover, although the WiX was developed based on the main existing theories of well-being, which are relevant internationally, the WiX so far has only been validated in the Netherlands. Future research should assess the validity of this new instrument in other countries, preferably also countries with different economic, political and cultural environments. Such studies would also provide opportunities for international comparisons of (determinants of) multi-dimensional subjective well-being. Direct comparison of the WiX to other instruments aiming to measure generic well-being, such as the ICECAP-A or the EQ-HWB (or, in older people, the ICECAP-O or the WOOP), are recommended to further explore the relative performance of instruments (in different subgroups and contexts). Finally, utility weights need to be determined for the items and response levels of the WiX to enable computation of utility scores for the well-being states described by the instrument. This would provide information about the relative importance of the items in a broader sample and enable the use of the WiX in economic evaluations of intervention in health and social care, but also in other sectors or across sectors.

Although further validation and valuation research is needed before the WiX can be applied in economic evaluations, for now we conclude that the WiX appears to be a promising alternative measure to assess well-being in the adult general population, able to capture the broad benefits of interventions in health and social care as well as other sectors.

References

Al-Janabi, H., Flynn, T.N., & Coast, J. (2012). Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. *Qual Life Res*, 21, 167-176.

Bartlett, M.S. (1951). The effect of standardization on a Chi-square approximation in factor analysis. *Biometrika*, 38, 337-344.

Brazier, J., Peasgood, T., Mukuria, C., Marten, O., Kreimeier, S., Luo, N., et al. (2022). The EQ-HWB: Overview of the Development of a Measure of Health and Wellbeing and Key Results. *Value Health*, 25, 482-491.

Brazier, J.E., Harper, R., Jones, N.M., O'Cathain, A., Thomas, K.J., Usherwood, T., et al. (1992). Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ*, 305, 160-164.

Cantril, H. (1965). *The Pattern of Human Concerns*. New Brunswick: Rutgers University Press.

Coast, J. (2014). Strategies for the economic evaluation of end-of-life care: making a case for the capability approach. *Expert Rev Pharmacoecon Outcomes Res*, 14, 473-482.

Coast, J., Peters, T.J., Natarajan, L., Sproston, K., & Flynn, T. (2008). An assessment of the construct validity of the descriptive system for the ICECAP capability measure for older people. *Qual Life Res*, 17, 967-976.

Davis, J.C., Bryan, S., Li, L.C., Best, J.R., Hsu, C.L., Gomez, C., et al. (2015). Mobility and cognition are associated with wellbeing and health related quality of life among older adults: a cross-sectional analysis of the Vancouver Falls Prevention Cohort. *BMC geriatr*, 15, 75.

Diener, E., Emmons, R.A., Larsen, R.J., & Griffin, S. (1985). The Satisfaction With Life Scale. *J Pers Assess*, 49, 71-75.

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, 94-122.

Drummond, M.F., Sculpher, M.J., Torrance, G.W., O'Brien, B.J., & Stoddart, G.L. (2005). *Methods for the Economic Evaluation of Health Care Programmes (3rd ed.)*. Oxford, UK: Oxford University Press.

EuroQol Group (1990). EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy*, 16, 199-208.

Furlong, W.J., Feeny, D.H., Torrance, G.W., & Barr, R.D. (2001). The Health Utilities Index (HUI) system for assessing health-related quality of life in clinical studies. *Ann Med*, 33, 375-384.

Grewal, I., Lewis, J., Flynn, T., Brown, J., Bond, J., & Coast, J. (2006). Developing attributes for a generic quality of life measure for older people: Preferences or capabilities? *Social Science & Medicine*, 62, 1891-1901.

Hackert, M.Q.N., Exel, J.V., & Brouwer, W.B.F. (2017). Valid Outcome Measures in Care for Older People: Comparing the ASCOT and the ICECAP-O. *Value Health*, 20, 936-944.

Hackert, M.Q.N., van Exel, J., & Brouwer, W.B.F. (2021). Content validation of the Well-being of Older People measure (WOOP). *Health and Quality of Life Outcomes*, 19, 200.

Hansen, T., & Slagsvold, B. (2012). The age and subjective well-being paradox revisited:A multidimensional perspective. *Norsk Epidemiologi*, 22.

Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., et al. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*, 20, 1727-1736.

Hopkins, W.G. (2002). *A Scale of Magnitudes for Effect Statistics: A New View of Statistics*.

Johnson, R., Jenkinson, D., Stinton, C., Taylor-Phillips, S., Madan, J., Stewart-Brown, S., et al. (2016). Where's WALY? : A proof of concept study of the 'wellbeing adjusted life year' using secondary analysis of cross-sectional survey data. *Health Qual Life Outcomes*, 14, 126.

Kaiser, H.F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 31-36.

Keeley, T., Coast, J., Nicholls, E., Foster, N.E., Jowett, S., & Al-Janabi, H. (2016). An analysis of the complementarity of ICECAP-A and EQ-5D-3 L in an adult population of patients with knee pain. *Health and Quality of Life Outcomes*, 14, 36.

Kennedy-Martin, M., Slaap, B., Herdman, M., van Reenen, M., Kennedy-Martin, T., Greiner, W., et al. (2020). 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.

Koo, T.K., & Li, M.Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*, 15, 155-163.

Krabbe, P.F., Stouthard, M.E., Essink-Bot, M.L., & Bonsel, G.J. (1999). The effect of adding a cognitive dimension to the EuroQol multiattribute health-status classification system. *J Clin Epidemiol*, 52, 293-301.

Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.

Makai, P., Brouwer, W.B., Koopmanschap, M.A., Stolk, E.A., & Nieboer, A.P. (2014). Quality of life instruments for economic evaluations in health and social care for older people: a systematic review. *Soc Sci Med*, 102, 83-93.

Mokkink, L.B., Prinsen, C.A.C., Patrick, D.L., Alonso, J., Bouter, L.M., De Vet, H.C.W., et al. (2018). COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). (p. 78).

Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., et al. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*, 63, 737-745.

Netten, A., Burge, P., Malley, J., Potoglou, D., Towers, A.M., Brazier, J., et al. (2012). Outcomes of social care for adults: developing a preference-weighted measure. *Health Technol Assess*, 16, 1-166.

Ulloa, B.F.L., Møller, V., & Sousa-Poza, A. (2013). How does subjective well-being evolve with age? A literature review. *Journal of Population Ageing*, 6.

Van der Deijl, W., Brouwer, W., & van Exel, J. (2023). What constitutes well-being? Five views among adult people from the Netherlands on what is important for a good life.

Versteegh, M.M., Vermeulen, K.M., Evers, S.M.A.A., de Wit, G.A., Prenger, R., & E, A.S. (2016). Dutch Tariff for the Five-Level Version of EQ-5D. *Value Health*, 19, 343-352.

Voormolen, D.C., Bom, J.A.M., de Bekker-Grob, E.W., Brouwer, W.B.F., & van Exel, J. (2023). Development and content validation of the 10-item Well-being instrument (WiX) for use in economic evaluation studies. *Submitted*.

Ware, J.E., Jr., & Sherbourne, C.D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*, 30, 473-483.

Wisløff, T., Hagen, G., Hamidi, V., Movik, E., Klemp, M., & Olsen, J.A. (2014). Estimating QALY gains in applied studies: a review of cost-utility analyses published in 2010. *Pharmacoeconomics*, 32, 367-375.

Appendix A. The 10-Item Well-being instrument (WiX) – English version

For each section, select the description that applies to you best today.

Mental health

Consider feeling mentally well and not suffering from feelings of anxiety, stress, sadness, and not having worries or a lack of control.

☐ I'm very satisfied with my mental health

☐ I'm satisfied with my mental health

☐ I'm not satisfied but also not dissatisfied with my mental health

☐ I'm dissatisfied with my mental health

☐ I'm very dissatisfied with my mental health

Physical health

Consider feeling physically well and not suffering from physical limitations, low energy, problems with vision, hearing, speech, sleep or movement, pain, and other physical complaints.

☐ I'm very satisfied with my physical health

☐ I'm satisfied with my physical health

☐ I'm not satisfied but also not dissatisfied with my physical health

☐ I'm dissatisfied with my physical health

☐ I'm very dissatisfied with my physical health

Relationships

Consider the relationships with your family, partner, friends, colleagues, and other people who are important to you. This concerns the number and quality of your relationships, including feeling love and friendship and getting practical or emotional support when you need it.

☐ I'm very satisfied with my relationships

☐ I'm satisfied with my relationships

☐ I'm not satisfied but also not dissatisfied with my relationships

☐ I'm dissatisfied with my relationships

☐ I'm very dissatisfied with my relationships

Living environment

Consider the house and the neighbourhood you live in, the availability of open and green areas, and the facilities and services that you need.

☐ I'm very satisfied with my living environment

☐ I'm satisfied with my living environment

☐ I'm not satisfied but also not dissatisfied with my living environment

☐ I'm dissatisfied with my living environment

☐ I'm very dissatisfied with my living environment

Safety

Consider feeling safe in your daily life, that others accept you and that you are not harassed because of who you are or what you think or believe, and that there is little risk of something terrible happening to you at home, on the streets or online.

☐ I'm very satisfied with my safety
☐ I'm satisfied with my safety
☐ I'm not satisfied but also not dissatisfied with my safety
☐ I'm dissatisfied with my safety
☐ I'm very dissatisfied with my safety

Financial situation

Consider having enough money for your expenses for things like housing, insurance, clothing, and food, occasionally something extra and not having any worries about money.

☐ I'm very satisfied with my financial situation
☐ I'm satisfied with my financial situation
☐ I'm not satisfied but also not dissatisfied with my financial situation
☐ I'm dissatisfied with my financial situation
☐ I'm very dissatisfied with my financial situation

Relaxation and leisure time

Consider taking a break, doing something fun, hobbies, sports and going on holiday. This concerns the amount and quality of your relaxation and leisure time.

☐ I'm very satisfied with my relaxation and leisure time
☐ I'm satisfied with my relaxation and leisure time
☐ I'm not satisfied but also not dissatisfied with my relaxation and leisure time
☐ I'm dissatisfied with my relaxation and leisure time
☐ I'm very dissatisfied with my relaxation and leisure time

Activities

Consider activities like studying, paid work, volunteer work, household chores and providing care or support to family, friends, or acquaintances. This concerns the amount and quality of your activities.

☐ I'm very satisfied with my activities
☐ I'm satisfied with my activities
☐ I'm not satisfied but also not dissatisfied with my activities
☐ I'm dissatisfied with my activities
☐ I'm very dissatisfied with my activities

Independence

Consider feeling autonomous and being able to find, understand and use information, and making your own choices and carrying them out in daily life.

☐ I'm very satisfied with my level of independence

☐ I'm satisfied with my level of independence

☐ I'm not satisfied but also not dissatisfied with my level of independence

☐ I'm dissatisfied with my level of independence

☐ I'm very dissatisfied with my level of independence

Self-worth

Consider living according to your own values and beliefs, and being satisfied with who you are and what you do.

☐ I'm very satisfied with my self-worth

☐ I'm satisfied with my self-worth

☐ I'm not satisfied but also not dissatisfied with my self-worth

☐ I'm dissatisfied with my self-worth

☐ I'm very dissatisfied with my self-worth

Appendix B

Table B1: Distribution (%) of responses to the WiX (N=1,045)

	Very dissatisfied	Dissatisfied	Not satisfied but also not dissatisfied	Satisfied	Very satisfied
Mental health	3.06	10.91	22.68	39.04	24.31
Physical health	4.98	16.17	24.98	38.09	15.79
Relationships	2.78	9.09	21.91	39.33	26.89
Living environment	2.11	5.45	17.61	48.42	26.41
Safety	1.15	4.31	17.8	50.81	25.93
Financial situation	5.93	12.25	25.65	37.61	18.56
Relaxation and leisure time	2.11	6.89	22.58	47.08	21.34
Activities	2.11	9.86	27.37	42.2	18.47
Independence	2.01	6.41	18.37	46.32	26.89
Self-worth	2.39	6.51	19.52	44.4	27.18

Table B2: OLS regression analyses of the items of the WiX on Satisfaction with Life Scale

		Satisfaction with Life Scale										
		Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**	**Model 9**	**Model 10**	**Model 11**
Mental health	Not satisfied, not dissatisfied	3.921***										1.282
	Satisfied	8.640***										3.380***
	Very satisfied	8.432***										3.066***
Physical	Not satisfied, not dissatisfied		3.639***									1.384*
health	Satisfied		6.614***									1.798**
	Very satisfied		5.945***									-0.0220
Relationships	Not satisfied, not dissatisfied			3.493***								0.964
	Satisfied			7.290***								2.122**
	Very satisfied			8.241***								2.669***
Living	Not satisfied, not dissatisfied				3.065***							0.735
environment	Satisfied				6.440***							1.608*
	Very satisfied				7.528***							1.911*
Safety	Not satisfied, not dissatisfied					4.397***						1.096
	Satisfied					7.215***						0.762
	Very satisfied					8.238***						0.622
Financial	Not satisfied, not dissatisfied						5.352***					2.962***
situation	Satisfied						8.313***					3.901***
	Very satisfied						8.270***					3.227***
Relaxation &	Not satisfied, not dissatisfied							4.503***				0.856
leisure time	Satisfied							8.266***				1.760*
	Very satisfied							9.166***				2.162*
Activities	Not satisfied, not dissatisfied								3.044***			0.216
	Satisfied								7.555***			1.620*
	Very satisfied								7.441***			-0.0799
Independ-ence	Not satisfied, not dissatisfied									3.255***		0.299

		Satisfaction with Life Scale										
		Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**	**Model 9**	**Model 10**	**Model 11**
	Satisfied									6.894***		0.0846
	Very satisfied									7.947***		0.642
Self-worth	Not satisfied, not dissatisfied										4.620***	1.320
	Satisfied										9.127***	2.014*
	Very satisfied										10.39***	3.024**
	Constant	10.56***	11.10***	10.38***	10.96***	9.834***	9.885***	10.19***	10.94***	11.47***	9.096***	2.925*
	N	1044	1044	1044	1044	1044	1044	1044	1044	1044	1044	1044
	R-squared	0.233	0.167	0.192	0.140	0.127	0.224	0.189	0.194	0.156	0.233	0.398

Notes: Including controls for age, sex, and education. The reference group for all variables is those individuals who are (very) dissatisfied. For the ease of interpretability, we assumed the outcome variable to be continuous and hence used OLS-regression estimates. * p<0.05, ** p<0.01, *** p<0.001

Table B3: OLS regression analyses of the items of the WiX of the Cantril Ladder

		Cantril ladder										
		Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**	**Model 9**	**Model 10**	**Model 11**
Mental health	Not satisfied, not dissatisfied	1.220***										0.562***
	Satisfied	2.000***										0.718***
	Very satisfied	2.570***										0.919***
Physical health	Not satisfied, not dissatisfied		1.019***									0.455***
	Satisfied		1.665***									0.636***
	Very satisfied		2.241***									0.630***
Relationships	Not satisfied, not dissatisfied			0.758***								0.038
	Satisfied			1.422***								0.129
	Very satisfied			1.989***								0.326
Living environment	Not satisfied, not dissatisfied				0.292							-0.306
	Satisfied				1.013***							-0.037
	Very satisfied				1.661***							0.019
Safety	Not satisfied, not dissatisfied					0.610*						-0.077
	Satisfied					1.303***						0.011
	Very satisfied					1.944***						0.122
Financial situation	Not satisfied, not dissatisfied						1.124***					0.574***
	Satisfied						1.772***					0.862***
	Very satisfied						2.625***					1.317***
Relaxation & leisure time	Not satisfied, not dissatisfied							1.225***				0.306
	Satisfied							1.840***				0.304
	Very satisfied							2.553***				0.452*
Activities	Not satisfied, not dissatisfied								1.169***			0.483**
	Satisfied								1.879***			0.577***
	Very satisfied								2.642***			0.729***
Independ-ence	Not satisfied, not dissatisfied									0.777***		-0.039

		Cantril ladder										
		Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**	**Model 9**	**Model 10**	**Model 11**
	Satisfied									1.546***		-0.005
	Very satisfied									1.961***		-0.117
Self-worth	Not satisfied, not dissatisfied										1.493***	0.731***
	Satisfied										2.341***	0.803***
	Very satisfied										2.828***	0.616**
	Constant	5.369***	5.374***	5.639***	6.011***	5.665***	5.306***	5.407***	5.280***	5.825***	4.985***	3.543***
	N	1044	1044	1044	1044	1044	1044	1044	1044	1044	1044	1044
	R-squared	0.253	0.214	0.171	0.133	0.134	0.259	0.194	0.232	0.147	0.243	0.426

Notes: Including controls for age, sex, and education. The reference group for all variables is those individuals who are (very) dissatisfied. For ease of interpretability, we assumed the outcome variable to be continuous and hence used OLS-regression estimates. * p<0.05, ** p<0.01, *** p<0.001.

Table B4: Spearman correlations of items of the WiX with items of EQ-5D-5L (N=1,045)

	Mobility	**Self-care**	**Usual activities**	**Pain and discomfort**	**Anxiety and depression**
Mental health	0.20***	0.15***	0.35***	0.33***	0.53***
Physical health	0.41***	0.21***	0.50***	0.57***	0.31***
Relationships	0.12***	0.12***	0.21***	0.21***	0.32***
Living environment	0.13***	0.14***	0.18***	0.18***	0.23***
Safety	0.17***	0.18***	0.22***	0.23***	0.27***
Financial situation	0.19***	0.10***	0.26***	0.31***	0.25***
Relaxation and leisure time	0.16***	0.16***	0.28***	0.25***	0.31***
Activities	0.22***	0.18***	0.35***	0.32***	0.35***
Independence	0.25***	0.25***	0.31***	0.27***	0.30***
Self-worth	0.20***	0.16***	0.32***	0.29***	0.44***

Table B5: Factor loadings for the items of the WiX and the domains of the EQ-5D-5L (N=1,045)

Variable	Factor1	Factor2	Factor3	Uniqueness
Mental health	0.704		0.409	0.292
Physical health	0.548	0.486		0.305
Relationships	0.658			0.543
Living environment	0.675			0.444
Safety	0.663			0.476
Financial situation	0.669			0.497
Relaxation and leisure time	0.738			0.398
Activities	0.730			0.402
Independence	0.679			0.432
Self-worth	0.780			0.328
Mobility		0.843		0.249
Self-care		0.764		0.275
Usual activities	0.335	0.825		0.187
Pain and discomfort	0.304	0.757		0.277
Anxiety & depression	0.469	0.399	0.397	0.369

Note: Oblimin rotation, loadings below 0.3 are dropped from the table to allow easier interpretation of results

Table B6: Factor loadings for the items of the WiX and the EQ-5D-5L + cognition bolt-on (N=1,045)

Variable	Factor1	Factor2	Factor3	Uniqueness
Mental health	0.599		0.314	0.288
Physical health	0.486	0.517		0.302
Relationships	0.652			0.542
Living environment	0.705			0.437
Safety	0.671			0.476
Financial situation	0.709			0.498
Relaxation and leisure time	0.743			0.397
Activities	0.716			0.402
Independence	0.635			0.432
Self-worth	0.719			0.327
Mobility		0.897		0.252
Self-care		0.672	0.336	0.261
Usual activities		0.775		0.186
Pain and discomfort		0.816		0.271
Anxiety & depression			0.648	0.339
Cognition		0.410	0.498	0.388

Note: Promax rotation, factor loadings below 0.3 are dropped from the table to allow easier interpretation of results

Table B7: Overview of differences in WiX scores by subgroups

Variables	N	WiX total score	Difference
Age			
18-24	116	37.17	-
25-34	187	38.59	
35-44	230	37.35	
45-54	228	36.89	
55-70	284	37.61	
Education			
Low	345	37.77	***
Mid	440	36.47	
High	260	38.95	
Make ends meet			
With (great) difficulty	507	35.68	***
(Fairly) easily	538	39.25	
Income			
Below €1500	200	34.60	***
€1500 and above	690	38.41	
Living situation			
Not living with a partner	456	36.70	***
Living with a partner	589	38.15	
Working status			
Working	672	38.86	***
Unemployed	58	33.66	
EQ-5D-5L			
EQ-5D-5L < mean	372	33.14	***
EQ-5D-5L > mean	673	39.94	
Cognition			
Average to severe cognition issues	115	31.99	***
No or some cognition issues	930	38.20	

Note: Education is categorized into Low (no, primary, pre-vocational education); Middle (secondary or middle vocational education); High (higher vocational or academic education)

Table B8: Overview of item scores of the WiX by subgroups

	N	WiX item score	Difference
Financial situation			
Making ends meet			
With (great) difficulty	507	3.04	***
(Fairly) easily	538	3.94	
Income			
Below €1500	200	2.79	***
€1500 and above	690	3.72	
Relations			
Marital status			
Not living with a partner	456	3.51	***
Living with a partner	589	4.00	
Mental health			
EQ-5D-5L			
EQ-5D-5L < mean	372	3.14	***
EQ-5D-5L > mean	673	4.02	
Physical health			
EQ-5D-5L			
EQ-5D-5L < mean	372	2.73	***
EQ-5D-5L > mean	673	3.83	
Independence			
Cognition			
Average to severe cognition issues	115	3.31	***
No or some cognition issues	930	3.97	

Abbreviations: EQ-5D-5L, five-levels EuroQol five-dimensional questionnaire

Table B9: Test-retest reliability estimates in retest sample (N=563)

	Correlation main and retest sample	Complete agreement	Weighted Kappa		
			mean	lower bound	upper bound
Mental health	0.715***	62.0%	0.73	0.67	0.78
Physical health	0.722***	58.1%	0.72	0.67	0.77
Relationships	0.648***	60.2%	0.64	0.58	0.71
Living environment	0.591***	64.8%	0.57	0.48	0.64
Safety	0.512***	62.7%	0.48	0.39	0.57
Financial situation	0.747***	65.2%	0.78	0.74	0.82
Relaxation and leisure time	0.579***	61.3%	0.59	0.52	0.66
Activities	0.470***	53.1%	0.49	0.41	0.56
Independence	0.614***	62.7%	0.62	0.55	0.68
Self-worth	0.641***	63.2%	0.69	0.62	0.75
WiX total	0.812***	15.3%	0.82[a]	0.79	0.85

Note: [a] For the total (continuous) WiX score two-way intra-class correlation is reported instead of weighted kappa.

Table B10: Test-retest reliability estimates in retest sample (N=316) – using linear weights

	Correlation main and retest sample	Complete agreement	Weighted Kappa		
			mean	lower bound	upper bound
Mental health	0.686***	61.1%	0.57	0.50	0.63
Physical health	0.704***	60.8%	0.58	0.51	0.64
Relationships	0.666***	63.3%	0.57	0.50	0.64
Living environment	0.630***	68.0%	0.53	0.45	0.61
Safety	0.519***	65.2%	0.44	0.35	0.53
Financial situation	0.775***	69.9%	0.68	0.62	0.74
Relaxation and leisure time	0.568***	63.0%	0.49	0.40	0.56
Activities	0.518***	57.0%	0.43	0.34	0.50
Independence	0.735***	69.0%	0.62	0.54	0.68
Self-worth	0.686***	64.6%	0.59	0.52	0.66
WiX total	0.809***	15.5%	0.82[a]	0.79	0.85

Note: [a] For the total (continuous) WiX score two-way intra-class correlation is reported instead of weighted kappa.

Erasmus University Rotterdam
Erasmus Centre for Health Economics Rotterdam
Burgemeester Oudlaan 50
3062 PA Rotterdam, The Netherlands
T +31 10 408 8555
E escher@eur.nl
W www.eur.nl/escher