



ScienceDirect

Contents lists available at sciencedirect.com
Journal homepage: www.elsevier.com/locate/jval

Themed Section

Construct Validity, Reliability, and Responsiveness of the 10-Item Well-being Instrument for Use in Economic Evaluation Studies

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

ABSTRACT

Objectives: Economic evaluations of interventions in health and social care require outcome measures that capture their full benefits, including those beyond health. This study aimed to assess construct validity, test-retest reliability, and responsiveness of the newly developed 10-item Well-being instrument (WiX).

Methods: Data were gathered via an online survey in a representative sample of the adult general population in The Netherlands (N = 1045). Construct validity was assessed by inspecting convergent, structural, and discriminant validity, following the COnsensus-based Standards for the selection of health status Measurement INstruments 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 2 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 WiX are favorable 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.

Keywords: measurement, validity, well-being, WiX.

VALUE HEALTH. 2024; ■(■):■-■

Highlights

- Economic evaluations of interventions in health and social care require outcome measures that capture their full benefits, including those beyond health. For this purpose, outcome measures capturing overall well-being comprehensively are needed.
- 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 comprehensively measures the satisfaction of members of the adult population on 10 important domains of wellbeing.
- The WiX seems to be a promising alternative for existing measures of well-being for evaluating interventions in health and social care.

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.¹ The benefits of interventions in the health domain are commonly captured in terms of quality-adjusted life-years, which comprise both length of life and health-related quality of life (HRQoL). Measures that focus on HRQoL may, however, fall short in capturing the full benefits of an intervention when not (only) physical or mental health but (also) impacts on broader aspects of quality of life are improved. Using incomplete information about the benefits of an intervention 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.²⁻⁴ The need for such broader measures may be even more evident in the context of prevention policies or

intersectoral interventions, aiming to improve health and other elements of well-being.

This need for broader outcome measures that can be used in economic evaluations has resulted in the introduction of several well-being instruments. One of these instruments is the 10-item Well-being instrument (WiX), which aims to measure well-being in the adult population 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.⁵ In contrast to existing measures that are grounded in capability theory,⁶⁻⁸ the WiX focuses on subjective well-being. Thus, the WiX assesses well-being based on individuals' judgments of their life situation, which include—but are not confined to—their capabilities and functionings. The scope of this well-being instrument was based on a theoretical framework synthesizing the main theories of well-being and an explorative study into what the general adult population in The Netherlands considers to be important constituents for a good life for themselves,⁹ alongside a scoping review of existing multidimensional instruments measuring satisfaction with life in

the adult population. A content validation study has shown that the WiX is relevant, comprehensive, and comprehensible.⁵ This study aimed to assess the construct validity, reliability, and responsiveness of the WiX, following the CONsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) methodology,¹⁰ in a representative sample of the adult population of The Netherlands.

Methods

Sampling Strategy

Data were gathered in 2 stages between October 7 to 12, 2021, and October 23 to 28, 2021, via 2 separate online surveys. For the first survey (main sample), a sampling agency recruited 1045 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. Two weeks after completing the first survey, respondents were invited to participate in a follow-up survey (retest sample) to test the reliability (in terms of test-retest) and responsiveness of the WiX. The time interval was chosen following COSMIN recommendations¹⁰ as being long enough to prevent recall bias while short enough to ensure limited individual changes in individual well-being. More than 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. Given that 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 10 domains.⁵ Each item offers 5 response levels, with score levels ranging from 1 “very dissatisfied” to 5 “very satisfied.” Aggregating these scores over all items, a total score for the WiX can be computed that ranges from 10 (lowest) to 50 (highest level of well-being). The English version of the instrument is presented in [Appendix A](#) in [Supplemental Materials](#) found at <https://doi.org/10.1016/j.jval.2024.02.014>. 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 2 surveys and, if so, to describe them.

In addition to the WiX, well-being was measured using the Satisfaction with Life Scale (SWLS), ranging from 5 (lowest) to 35 (highest satisfaction with life),¹¹ and the Cantril Ladder, ranging from 0 (worst) to 10 (best possible life).¹² Both measures have a similar focus as the WiX on experienced well-being.

The health of respondents was measured using the EQ-5D instrument (EQ-5D-5L) and the EQ-visual analog scale (EQ-VAS). The EQ-5D-5L¹³ aims to measure and value HRQoL on 5 dimensions. Utility weights for The Netherlands were used to calculate utility scores,¹⁴ with 1 representing perfect health and 0 representing dead. A cognition bolt-on for the EQ-5D-5L was added to the survey to capture issues with concentration and memory.^{15,16} Furthermore, respondents were asked to rate their overall health today on the EQ-VAS, ranging from 0 (worst) to 100 (best imaginable health).¹³

Background Characteristics

The following background characteristics were included in the first survey: age; sex; household composition; educational

attainment, classified as low (no, primary, prevocational 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 4 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.¹⁷ 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 of the total sample (equaling a completion time for the full survey of <87 seconds). This did not substantially affect the findings from the analyses or their interpretation.

Descriptive statistics were computed for all measures. In addition, correlations among WiX items were inspected to assess their relevance. Given that the WiX was designed as a multidimensional measure of well-being, the items were expected to be positively correlated but not highly, because they supposedly represent distinct domains of well-being. Very high correlation between items could indicate redundancy. Following guidelines of Hopkins (2002),¹⁸ strength of correlation was evaluated as follows: <0.10, trivial; 0.10 to 0.29, small; 0.30 to 0.49, moderate; 0.50 to 0.69, high; 0.70 to 0.89, very high; and ≥ 0.90 , (nearly) perfect.

Construct validity encompasses convergent, structural, and discriminative/known-groups validity. Convergent validity concerns the degree to which an instrument 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 2 alternative measures of experienced well-being (Cantril Ladder and 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. Finally, we assessed the correlation between the WiX and 2 health measures (EQ-VAS and EQ-5D-5L). Here we expected high correlations of the scores on the mental and physical health items, with scores on these measures, and lower but still positive correlations for the other WiX items.

Structural validity concerns whether the scores of an instrument adequately reflect the dimensionality of the construct. This was analyzed using exploratory factor analysis (EFA). In particular, we assessed the overlap in factors between the WiX and the EQ-5D-5L. We expected the items of the WiX to correspond with a larger set of factors than the items of the EQ-5D-5L given that it measures broader outcomes. Bartlett's test of sphericity¹⁹ was used to ensure that the correlation matrix was not random, and the Kaiser-Meyer-Olkin measure of sampling adequacy²⁰ was applied to inspect whether the data were suited for factor analysis. Given that the items of both measures concern ordinal variables, we calculated a polychoric correlation matrix to use for the principal factor analysis. The appropriate number of factors was selected based on the Kaiser criterion, the scree plot, and the interpretability of the models. To allow the factors to be correlated, oblique rotation was applied. Promax was used for the main analysis; oblimin was used as a robustness check. Furthermore, as a robustness check, the analysis was repeated including the cognition bolt-on next to the original EQ-5D-5L.

Discriminative or known-groups validity concerns whether the instrument can discriminate between relevant subgroups. This was investigated by inspecting whether scores on the WiX (items) differed between subgroups as expected, using *t* tests and one-way analyses of variance. Based on previous research, we expected respondents of higher age, with higher level of education, who are employed, with higher income, who are able to make ends meet (fairly) easily, and in a relationship to report higher levels of well-being.²¹ In addition, for individuals with financial difficulties, we expected lower scores for the item “Financial situation,” whereas for individuals in a relationship we expected higher scores on the item “Relations.” Moreover, previous literature showed strong associations between health and well-being.²¹ Hence, we expected individuals in poor health based on their EQ-5D-5L scores to have lower total WiX scores and lower scores on the items “Mental health” and “Physical health.” Because cognition was found to be an important element of well-being,²² we expected lower WiX scores among those reporting issues with cognition.

The reliability and responsiveness of the WiX were assessed using the data from the retest sample. To evaluate test-retest reliability, 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 reporting 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, applying the following interpretation of the scores: 0.21 to 0.40, fair; 0.41 to 0.60 moderate; 0.61 to 0.80 substantial; and 0.81 to 1.00, (nearly) perfect.²³ The consistency in responses on total WiX scores was evaluated by calculating intraclass correlation using a 2-way mixed effects model,²⁴ which were interpreted as follows: 0.5 to 0.75, moderate; 0.75 to 0.90, good; and >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 2-week time interval between 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 reported a positive or negative event and those who did not.

Ethics

The study was approved by the Research Ethics Review Committee of the Erasmus School of Health Policy & Management (case number 21-001). All respondents provided an informed consent.

Results

Sample Characteristics

Descriptive characteristics of our main sample and the retest sample are presented 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.

The distribution of scores on the items of the WiX is presented in Figure 1 (the underlying scores are presented in Appendix Table B1 in Supplemental Materials found at <https://doi.org/10.1016/j.jval.2024.02.014>).

Table 1. Descriptive statistics main sample and retest sample.

Variables	Main sample (N = 1045), %	Retest sample (n = 563), %	Difference
Sex			
Male	50.1	43.9	<i>P</i> < .05
Female	49.8	56.1	
Other/prefer not to tell	0.1	-	
Age			
18-24	11.1	6.9	<i>P</i> < .001
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. Education is categorized into low (no, primary, prevocational education), middle (secondary or middle vocational education), high (higher vocational or academic education).

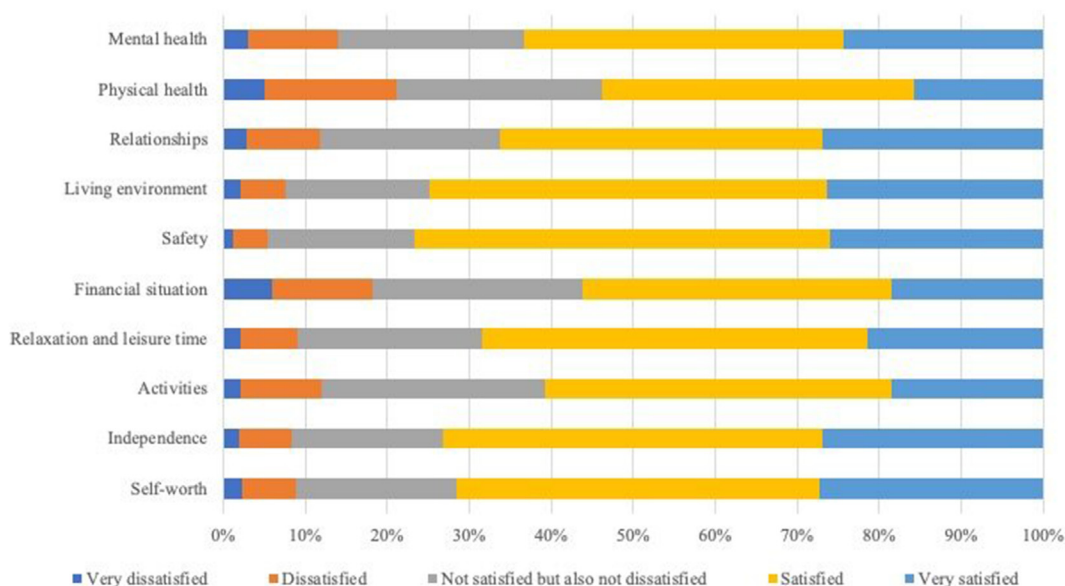
016/j.jval.2024.02.014). Most respondents indicated to be (very) satisfied on all items, with the lowest proportion of individuals indicating to be (very) satisfied with their physical health (53.9%) and their financial situation (56.1%). Few individuals were (very) dissatisfied with their safety (5.5%) or living environment (7.6%). The resulting mean total WiX score in the main sample was 37.5 (SD 6.84); the distribution of these scores is presented in Appendix Figure 1 in Supplemental Materials found at <https://doi.org/10.1016/j.jval.2024.02.014>. Although a large proportion of respondents reported to be (very) satisfied on most of the items, the ceiling effect seems to be moderate for total WiX scores: 5.2% reported to be very satisfied on all items of the WiX. Only 1 respondent indicated to be very dissatisfied on all items of the WiX.

The correlations between scores on the items of the WiX are presented in Table 2. Most items were moderately correlated, supporting the relevance of the separate items, but some items were highly correlated.

Convergent Validity

The correlations of (the items of) the WiX with the Cantril Ladder and SWLS were moderate to high (see Appendix Table B2 in Supplemental Materials found at <https://doi.org/10.1016/j.jval.2024.02.014>). Regression analysis of the items of the WiX on 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 (Appendix Tables B3 and B4 in Supplemental Materials found at <https://doi.org/10.1016/j.jval.2024.02.014>). When adding all items in 1 model, nearly all items were still positively associated with scores on both measures but not all associations were statistically significant.

In addition, correlations of scores on (the items of) the WiX with the EQ-5D-5L and EQ-VAS showed high correlations for the health-related items of the WiX, especially for “Physical health” (Appendix Table B2 in Supplemental Materials found at <https://doi.org/10.1016/j.jval.2024.02.014>).

Figure 1. Distribution of responses on the 10 items of the WiX (N = 1045).

WiX indicates 10-item Well-being instrument.

doi.org/10.1016/j.jval.2024.02.014). Correlations with other items of the WiX, such as “Living environment” and “Safety,” were small. Correlations between the items of the WiX and the items of the EQ-5D-5L are presented in [Appendix Table B5](https://doi.org/10.1016/j.jval.2024.02.014) (in [Supplemental Materials](https://doi.org/10.1016/j.jval.2024.02.014) found at <https://doi.org/10.1016/j.jval.2024.02.014>); correlations were especially high 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.”

Structural Validity

Using EFA to assess the overlap in factors between the WiX and the EQ-5D-5L, we identified 3 factors ([Table 3](#)). The first factor contained only items of the WiX, whereas the other 2 were a

combination of both WiX and EQ-5D-5L items. Factor 2 seems to capture items related to mental health and factor 3 items related to physical health.

The results of 2 alternative models, which show slightly different results but also clearly indicate that the WiX seems to capture dimensions of well-being beyond health, are presented in [Appendix Tables B6 and B7](#) (in [Supplemental Materials](https://doi.org/10.1016/j.jval.2024.02.014) found at <https://doi.org/10.1016/j.jval.2024.02.014>).

Discriminative/Known-Groups Validity

Mean WiX scores for subgroups in the sample are presented in [Figure 2](#). Except for age, we found scores to differ as expected between subgroups. Significantly higher WiX (item) scores were

Table 2. Spearman correlations of items of the WiX (N = 1045).

WiX items	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

Note. All correlations are significant at $P < .001$.
WiX indicates 10-item Well-being instrument.

Table 3. Factor loadings for the items of the WiX and the EQ-5D-5L (N = 1045).

Variable	Factor 1	Factor 2	Factor 3	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 and depression		0.641		0.369

Note. Promax rotation, factor loadings below 0.3 are dropped from the table to allow easy interpretation of results. WiX indicates 10-item Well-being instrument.

observed among respondents with higher education, making ends meet (fairly) easily, living together with a partner, being employed, and being in better health (Appendix Tables B8 and B9 in Supplemental Materials found at <https://doi.org/10.1016/j.jval.2024.02.014> present the underlying estimates).

Reliability

We assessed test-retest reliability when limiting our sample to respondents who indicated not to have experienced an event that affected their well-being significantly in the 2-week time interval between surveys (n = 316; 56.1%) (see Appendix Table B10 in Supplemental Materials found at <https://doi.org/10.1016/j.jval.2024.02.014>). The correlations between WiX (item) scores in the main and retest samples were mostly high to very high. The percentages complete agreement 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, indicating moderate to substantial agreement,²³ whereas the intraclass correlation score of 0.82 for the total WiX scores can be interpreted as good reliability.²⁴ The results for the full retest sample (n = 563) and when using linear weights (presented in Appendix Tables B11 and B12 in Supplemental Materials found at <https://doi.org/10.1016/j.jval.2024.02.014>) were very similar.

Responsiveness

In the retest sample, 247 respondents (43.9%) reported an event that affected their well-being significantly in the 2 weeks between surveys. Among them, 115 respondents reported having experienced an event that positively affected their well-being, 74 respondents reported a negative event, and 58 respondents reported both. Respondents who experienced a positive event most often mentioned having been on holidays or having experienced something pleasant such as meeting a friend/family member or going out for dinner, whereas those who experienced a negative event often mentioned something related to their mental or physical health. Compared with those not experiencing a significant event within the 2-week interval, experiencing a positive event was associated with a statistically significant increase in the

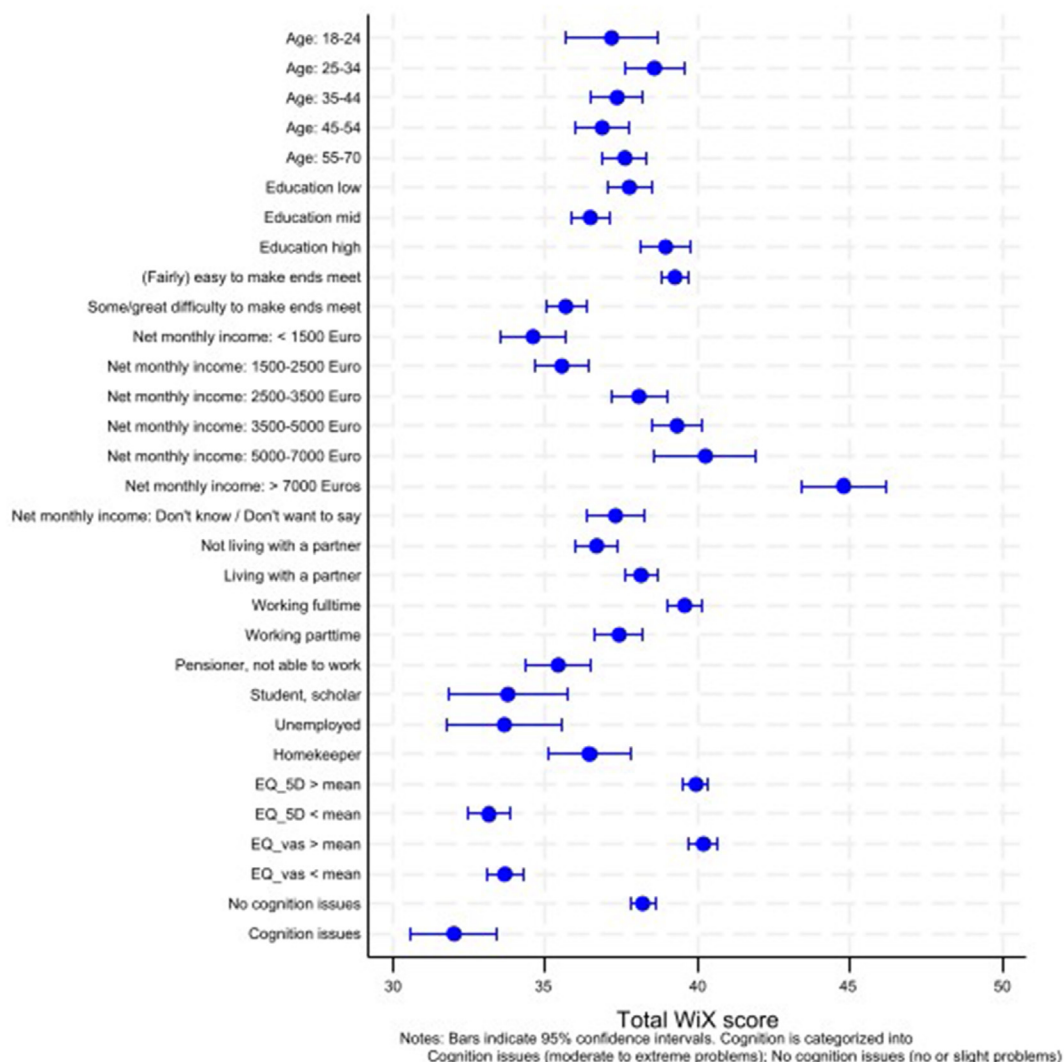
total WiX score of 1.24 (95% CI 0.31–2.17). Experiencing a negative event was not associated with a change in WiX score.

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 new instrument to comprehensively measure the satisfaction of adults on 10 important domains of well-being. The analyses demonstrated that the items of the WiX seem to be relevant and the construct validity and test-retest validity of the instrument seem to be sufficient.

Correlations between the WiX (items) and the SWLS and the Cantril Ladder were positive and in nearly all cases moderate to 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. Furthermore, the variation in correlations of WiX items with the 2 alternative subjective well-being measures may partly reflect differences in the concepts these measures capture. To further assess the relevance of specific WiX items in measuring well-being, future research should focus on studying subgroups of individuals who experience lower levels of well-being and are expected to show more variation in scores on the WiX items.

Using EFA, the items of the WiX showed to be related with items of the EQ-5D-5L in the expected way, which indicates structural validity. The dimensionality analysis identified 2 factors capturing the health-related items of the WiX and the domains of the EQ-5D-5L, representing the mental and physical health dimensions of well-being, and a third factor comprising the nonhealth items of the WiX. The latter can be interpreted as elements considered important for well-being that are not covered by HRQoL measures

Figure 2. Overview of average total WiX scores by subgroups.

WiX indicates 10-item Well-being instrument.

(and possibly also not by the EQ Health and Well-being, given that it was developed to measure health-related well-being²⁵). These results align with earlier studies demonstrating that well-being measures capture elements beyond health, not captured by HRQoL measures such as the EQ-5D-5L.^{26,27} The results of the dimensionality analysis also emphasize that, as intended, the WiX is a comprehensive measure capturing both (mental and physical) health and broader elements of well-being.

When comparing WiX scores between subgroups, the observed differences were largely in line with expectations based on previous research.²¹ Individuals with a higher education and 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 among individuals in different age groups. Potentially, this relationship observed in the literature is driven by other elements associated with old age, such as a decline in health.²⁸ In addition, the hypothesized increase in well-being might mainly occur in individuals aged 70 years and older,²⁹ who were not included in our sample. The timing of our study (ie, during the COVID-19 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 a less stable nature of items such as “Relaxation and leisure time” compared with health and finances when measured over a 2-week interval. Earlier research indicated that physical constructs often show to be more reliable than social ones.³⁰

Evidence regarding the responsiveness of the WiX is still limited. Only a small proportion of the retest sample mentioned to have experienced an event significantly affecting their well-being in the 2 weeks between surveys. Among them, we found a significant increase in WiX score for those experiencing positive events, but not a significant change for negative events. These findings may relate to the type of events reported by respondents, which mostly concerned relatively events such as a visit to the theater or having a cold.

Future (intervention) studies in larger samples experiencing potentially more impactful events are needed to assess the responsiveness of the WiX adequately. Another important aspect in relation to the validity of the WiX for policy evaluation concerns the role of adaptation.^{31,32} The current study cannot provide evidence about the extent to which scores are influenced by individuals adjusting to lower states on certain dimensions of well-being over time and whether this is more prevalent in certain subgroups (eg, older people adapting to deteriorating health).

Although 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. Although ceiling effects seem moderate when examining the overall WiX score, for several items there is clustering of responses toward the upper end of the response scale. This may be attributed to the high levels of life satisfaction generally reported in The Netherlands.³³ In addition, the timing of our study (ie, during the COVID-19 pandemic) and the use of an online survey should be considered in interpreting our findings. It seems relevant to further explore the psychometric properties of the WiX 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 instrument in other populations, preferably in countries with different economic, political, and cultural environments. Such studies would also provide opportunities for international comparisons of (determinants of) multidimensional subjective well-being. Direct comparison of the WiX to other instruments aiming to measure well-being, such as the ICEpop Capability Measure for Adults or the EQ Health and Well-being, is recommended to further explore the relative performance of these different instruments (in different subgroups and contexts). Finally, utility weights need to be determined for 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. For the latter, it is important to bear in mind that intersectoral comparisons may be difficult to undertake and act upon in both policy and practice and hence need careful consideration.

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

Author Disclosures

Links to the disclosure forms provided by the authors are available [here](#).

Supplemental Material

Supplementary data associated with this article can be found in the online version at <https://doi.org/10.1016/j.jval.2024.02.014>.

Article and Author Information

Accepted for Publication: February 25, 2024

Published Online: xxxx

doi: <https://doi.org/10.1016/j.jval.2024.02.014>

Author Affiliations: Erasmus School of Health Policy & Management, Erasmus University Rotterdam, Rotterdam, The Netherlands (Bom, Voormolen, Brouwer, de Bekker-Grob, van Exel); Erasmus Centre for Health Economics Rotterdam (EsCHER), Erasmus University Rotterdam, Rotterdam, The Netherlands (Bom, Voormolen, Brouwer, de Bekker-Grob, van Exel); Erasmus Choice Modelling Centre, Erasmus University Rotterdam, Rotterdam, The Netherlands (de Bekker-Grob, van Exel).

Correspondence: Daphne C. Voormolen, PhD, Erasmus School of Health Policy & Management, Erasmus University Rotterdam, Burgemeester Oudlaan 50, P.O. Box 1738, 3000 DR, Rotterdam, The Netherlands. Email: voormolen@eshpm.eur.nl

Author Contributions: *Concept and design:* Bom, Voormolen, Brouwer, de Bekker-Grob, van Exel

Acquisition of data: Bom, Voormolen

Analysis and interpretation of data: Bom, Voormolen

Drafting of the manuscript: Bom, Voormolen

Critical revision of the paper for important intellectual content: Bom, Voormolen, Brouwer, de Bekker-Grob, van Exel

Statistical analysis: Bom, Voormolen

Obtaining funding: de Bekker-Grob

Supervision: Brouwer, de Bekker-Grob, van Exel

Funding/Support: The authors received no financial support for this research.

Acknowledgment: Erasmus Initiative (Smarter Choices for Better Health).

REFERENCES

1. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the Economic Evaluation of Health Care Programmes*. 3rd ed. Oxford, United Kingdom: Oxford University Press; 2005.
2. Makai P, Brouwer WB, Koopmanschap MA, Stolk EA, Nieboer AP. Quality of life instruments for economic evaluations in health and social care for older people: a systematic review. *Soc Sci Med*. 2014;102:83–93.
3. Johnson R, Jenkinson D, Stinton C, et al. 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*. 2016;14(1):126.
4. Coast J. Strategies for the economic evaluation of end-of-life care: making a case for the capability approach. *Expert Rev Pharmacoecon Outcomes Res*. 2014;14(4):473–482.
5. Voormolen DC, Bom JAM, de Bekker-Grob EW, Brouwer WBF, van Exel J. Development and content validation of the 10-item Well-being instrument (WiX) for use in economic evaluation studies. *Appl Res Qual Life*. 2024;19:381–413.
6. Sen A. *Commodities and Capabilities*. Amsterdam, The Netherlands: North-Holland; 1985.
7. Binder M. Subjective well-being capabilities: bridging the gap between the capability approach and subjective well-being research. *J Happiness Stud Interdiscip Forum Subjective Well-Being*. 2014;15(5):1197–1217.
8. Robeyns I. The Capability Approach: a theoretical survey. *J Hum Dev*. 2005;6(1):93–117.
9. van der Deijl W, Brouwer W, van Exel J. What constitutes well-being? Five views among adult people from the Netherlands on what is important for a good life. *Appl Res Qual Life*. 2023;18(6):3141–3167.
10. Mokkink LB, Prinsen CAC, Patrick DL, et al. *COSMIN Methodology for Systematic Reviews of Patient-Reported Outcome Measures (PROMs)*. 1. 2018:78. https://efaidnbmnnnibpcajpcgiclfndmkaj/https://cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018.pdf. Accessed March 23, 2024.
11. Diener E, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. *J Pers Assess*. 1985;49(1):71–75.
12. Cantril H. *The Pattern of Human Concerns*. New Brunswick, NJ: Rutgers University Press; 1965.
13. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res Int J Qual Life Aspects Treat Care Rehabil*. 2011;20(10):1727–1736.
14. Versteegh MM, Vermeulen KM, Evers SMAA, de Wit GA, Prenger R, E AS. Dutch tariff for the five-level version of EQ-5D. *Value Health*. 2016;19(4):343–352.

15. Krabbe PF, Stouthard ME, Essink-Bot ML, Bonsel GJ. The effect of adding a cognitive dimension to the EuroQol multiattribute health-status classification system. *J Clin Epidemiol*. 1999;52(4):293–301.
16. Finch AP, Brazier JE, Mukuria C. Selecting bolt-on dimensions for the EQ-5D: examining their contribution to health-related quality of life. *Value Health*. 2019;22(1):50–61.
17. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737–745.
18. Hopkins WG. *A Scale of Magnitudes for Effect Statistics: A New View of Statistics*. 502. 2002:411.
19. Bartlett MS. The effect of standardization on a chi-square approximation in factor analysis. *Biometrika*. 1951;38(3-4):337–344.
20. Kaiser HF. An index of factorial simplicity. *Psychometrika*. 1974;39(1):31–36.
21. Dolan P, Peasgood T, White M. Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *J Econ Psychol*. 2008;29(1):94–122.
22. Davis JC, Bryan S, Li LC, et al. 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*. 2015;15(1):75.
23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174.
24. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–163.
25. Brazier J, Peasgood T, Mukuria C, et al. The EQ-HWB: overview of the development of a measure of health and wellbeing and key results. *Value Health*. 2022;25(4):482–491.
26. Hackert MQN, van Exel J, Brouwer WBF. Content validation of the Well-being of Older People measure (WOOP). *Health Qual Life Outcomes*. 2021;19(1):200.
27. Keeley T, Coast J, Nicholls E, Foster NE, Jowett S, Al-Janabi H. An analysis of the complementarity of ICECAP-A and EQ-5D-3 L in an adult population of patients with knee pain. *Health Qual Life Outcomes*. 2016;14(1):36.
28. Ulloa BFL, Møller V, Sousa-Poza A. How does subjective well-being evolve with age? A literature review. *J Popul Ageing*. 2013;6(3):227–246.
29. Hansen T, Slagsvold B. The age and subjective well-being paradox revisited: A multidimensional perspective. *Norsk Epidemiol*. 2012;11(29):22.
30. Brazier JE, Harper R, Jones NM, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ*. 1992;305(6846):160–164.
31. Frederick S, Loewenstein G. Hedonic adaptation. In: Kahneman D, Diener E, Schwarz N, eds. *Well-Being: The Foundations of Hedonic Psychology*. New Brunswick, NJ: Russell Sage Foundation; 1999:302–329.
32. Stöckel J, van Exel J, Brouwer WBF. Adaptation in life satisfaction and self-assessed health to disability - Evidence from the UK. *Soc Sci Med*. 2023;328:115996.
33. Netherlands. OECD Better Life Index. <https://www.oecdbetterlifeindex.org/countries/netherlands/>. Accessed March 23, 2024.