FORS

explore.understand.share.

Carmen Borrat-Besson, Valérie-Anne Ryser and Judite Gonçalves

An evaluation of the CASP-12 scale used in the Survey of Health, Ageing and Retirement in Europe (SHARE) to measure Quality of Life among people aged 50+

FORS Working Papers 2015-4

FORS Working Paper series

The FORS Working Paper series presents findings related to survey research, focusing on methodological aspects of survey research or substantive research. Manuscripts submitted are papers that represent work-in-progress. This series is intended to provide an early and relatively fast means of publication prior to further development of the work. A revised version might be requested from the author directly.

Further information on the FORS Working Paper Series can be found on www.forscenter.ch

How to cite this document:
Borrat-Besson, C., Ryser, V.-A., & Gonçalves, J. (2015). An evaluation of the CASP-12 scale used in the Survey of Ageing and Retirement in Europe (SHARE) to measure Quality of Life among people aged 50+. *FORS Working Paper Series,* paper 2015-4. Lausanne: FORS.

Acknowledgments:
This paper uses data from SHARE Wave 5 release 1.0.0, as of March 31st 2015 (DOI: 10.6103/SHARE.w5.100) or SHARE Wave 4 release 1.1.1, as of March 28th 2013 (DOI: 10.6103/SHARE.w4.111) or SHARE Waves 1 and 2 release 2.6.0, as of November 29th 2013 (DOI: 10.6103/SHARE.w1.260 and 10.6103/SHARE.w2.260) or SHARELIFE release 1.0.0, as of November 24th 2010 (DOI: 10.6103/SHARE.w3.100). The SHARE data collection has been primarily funded by the European Commission through the 5th Framework Programme (project QLK6-CT-2001-00360 in the thematic programme Quality of Life), through the 6th Framework Programme (projects SHARE-I3, RII-CT-2006-062193, COMPARE, CIT5- CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th Framework Programme (SHARE-PREP, N° 211909, SHARE-LEAP, N° 227822 and SHARE M4, N° 261982). Additional funding from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, R21 AG025169, Y1-AG-4553-01, IAG BSR06-11 and OGHA 04-064) and the German Ministry of Education and Research as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions).

ISSN 1663-523x (online)

FORS — c/o University of Lausanne, Géopolis — 1015 Lausanne, Switzerland
E-mail: paperseries@fors.unil.ch

This article evaluates the CASP-12 scale used in the Survey of Health, Aging and Retirement in Europe (SHARE). The CASP-12 is a shorter version of the CASP-19, a measure of quality of life (QoL) in older ages. The CASP-19 is based on a sociological conceptualization of QoL that draws upon the "Theory of Human Need" and has four dimensions: Control, Autonomy, Self-realization and Pleasure.

We evaluate the structure of the SHARE version of the CASP scale using internal consistency analyses, factor analyses and item-total Spearman correlations. In addition, we assess whether that structure is invariant across the various countries participating in SHARE. Finally, we propose and test a revised scale based on our results.

The structure postulated by the authors of the CASP-19 could not be replicated with the CASP-12 scale used in SHARE. Factor analyses results suggest a revised scale with ten items instead of twelve, and two factors instead of four. Cross-country comparisons showed that results were similar in all but two countries, Italy and Portugal.

An evaluation of the CASP-12 scale used in the Survey of Health, Ageing and Retirement in Europe (SHARE) to measure Quality of Life among people aged 50+

Carmen Borrat-Besson[1], Valérie-Anne Ryser[2], & Judite Gonçalves[3]

1. Introduction

In the last decades, individuals' quality of life (QoL) has become a major and unavoidable social issue, with health actors and policy makers developing different policy interventions aiming to improve individuals' QoL. The attention is often focused on older individuals, since they are most likely to experience negative events that jeopardize their autonomy and more generally the quality of their everyday life (e.g. hospitalization, institutionalization, disease, death of friends and family members). However, the assessment of people's QoL faces important theoretical and methodological challenges. What is high or low quality of life? How can we measure the quality of people's lives?

Reviewing the literature related to the QoL of elderly people, Higgs and colleagues note that many measurement instruments are not rooted in explicit and well-defined theoretical frameworks (Higgs et al., 2003; Hyde et al., 2003). Scales are often developed ad hoc and according to the needs of the different studies. In response to this lack of theory-grounded measures, the same authors propose a measure of QoL for individuals aged 65 to 75, the so-called CASP-19, based on the needs-satisfaction theory (Maslow, 1943; Doyal & Gough, 1991). This scale includes 19 Likert-type items reflecting four different dimensions of QoL: Control, Autonomy, Self-realization, and Pleasure. The first letter of each dimension form the acronym CASP.

[1] FORS – Lausanne, carmen.borrat-besson@fors.unil.ch.
[2] FORS – Lausanne
[3] Geneva School of Economics and Management, University of Geneva

Several studies, principally conducted in the United Kingdom, tested the psychometric properties of the CASP-19 using exploratory and confirmatory factor analysis (Bowling, 2009; Bowling & Stenner, 2011; Howel, 2012; Sim et al., 2011; Wiggins et al., 2008; Wu et al., 2013). A shorter version with only 12 items, the so-called CASP-12, was also proposed and tested (Wiggins et al., 2008). The Survey of Health, Ageing and Retirement in Europe (SHARE) introduced a 12-item version of the CASP scale that differs from the CASP-12 suggested by Wiggins and colleagues (2008) and has not yet been tested (von dem Knesebeck, Hyde, Higgs, Kupfer, & Siegriest, 2005).

The aim of this study is twofold: first, to evaluate the psychometric properties of the SHARE version of the CASP scale, and second, to explore its cross-cultural robustness. Three reasons motivate our research. First, the SHARE version of the CASP scale has not been thoroughly evaluated and there is no evidence that it will present the same psychometric properties as the CASP-12 suggested by Wiggins and colleagues (2008). Second, the age range of the SHARE participants (50+) is much broader than the age range of the target population for which the CASP scale was developed (65-75). There is no evidence that the CASP scale is a good measure of QoL of this broader population. And third, the SHARE version of the CASP scale was translated and applied in all participating countries (Austria, Belgium, Czech Republic, Denmark, Estonia, France, Germany, Hungary, Italy, Netherlands, Poland, Portugal, Slovenia, Spain, Sweden and Switzerland) without evidence that it can be exported to different cultural settings.

In the following sections, we first discuss the theoretical and empirical background of the CASP-19 and the CASP-12. Second, we present the SHARE dataset on which our study is based and report some descriptive statistics. Third, we report the results of two analyses examining the psychometric properties of the SHARE version of the CASP-12. In the first analysis, we test the factor structure postulated by the theory. In the second analysis, we test a modified version of the CASP that includes only 10 items. The final section draws conclusions and discusses further implications and recommendations for using the SHARE version of the CASP-12.

2. Theoretical background

The CASP-19 was developed to measure the QoL of middle-aged individuals (65-75), with three main objectives (Higgs et al., 2003; Hyde et al., 2003). First, the authors wanted to derive an operational definition of QoL from an explicit and well-defined theoretical framework. Second, they wanted to provide a measure of QoL that didn't use the

determinants of QoL as indicators of QoL. For example, health indicators are frequently used as proxies for QoL. However, health is a factor which may influence people's QoL and should therefore not be used as an indicator of QoL. In using health as a proxy for QoL, researchers rely on strong normative assumptions, namely that QoL is high when people are in good health and low when they are in bad health. However, people – especially very old people – have efficient strategies to adapt to health deterioration and more generally losses associated with old and very old age. Therefore, older persons may often report good quality of life despite the decline in health. Third, the authors' approach aimed at providing a measure of QoL that allows a meaningful comparison between individuals. According to the authors, measures of QoL that rely on subjective evaluations (e.g. life satisfaction) may suffer from judgmental relativity, because the standards people refer to when assessing their lives differ from one person to another. Therefore, the evaluations that people give may not be comparable. To provide a meaningful comparison, the measurement instrument has to rely on fundamental principles that are true for all individuals.

To fulfill these three objectives, Higgs and colleagues draw upon the "Theory of Human Need". This theoretical framework goes back to Maslow's hierarchy of human needs, according to which individuals have an intrinsic motivation to fulfill a common set of needs (Maslow, 1943). These needs are rooted in human nature and meeting them is an indicator of good psychological functioning, personal growth, and well-being. They are supposed to be universal and objective. The authors assume that measuring the extent to which needs are fulfilled provides a measure of QoL that is more objective than a personal evaluation (e.g. life satisfaction), therefore allowing for "a meaningful comparison between people's different QoL scores" (Higgs et al., 2003, p. 244).

Based on Maslow (1943) and Doyal and Gough (1991), Higgs and colleagues identify four dimensions of needs from which they derive the CASP items (Higgs et al., 2003; Hyde et al., 2003). The four dimensions are Control, Autonomy, Self-realization, and Pleasure. Control is defined as the perception of being able to shape one's own life, to have control over one's environment through one's own behaviors (Rotter & Mulry, 1965). Autonomy refers to self-determination and the absence of unwanted interference from others. In the words of Deci and Ryan (1987), the concept of autonomy refers to "an inner endorsement of one's actions, the sense that they emanate from oneself and are one's own". Pleasure refers to the pursuit of enjoyable activities. And finally, Self-realization describes the fulfillment of oneself. According to Higgs and colleagues (2003), these two last dimensions measure a new feature of the young-old individuals born after World War II. This generation benefited from favorable life conditions (e.g. progress in medicine,

expansion of retirement systems), that improved their general health conditions and contributed to the increase in life expectancy. They reach, in general, retirement age with several healthy years ahead of them. Relieved from professional responsibilities, they have time to pursue new interests and enjoy life. Retirement age is seen by the authors as a period for self-realization, because people have time to invest new interests or develop sleeping interests they always neglected during their professional life. They can flourish in activities they enjoy and that correspond best to their personality. Whereas Maslow's approach assumes a hierarchical organization of needs, from the most basic to increasingly complex ones, Wiggins and colleagues assume, like Doyal and Dough, that all needs have the same importance. The original CASP scale comprises 19 Likert-type items; four to five items per dimension (Table 1). These items can be combined into a measure of overall QoL.

Table 1: List of the CASP items by dimension, used in the different surveys

Dimensions	Items	CASP-19 (Hyde et al., 2003)	CASP-12 (Wiggins et al., 2008)	SHARE version oft he CASP-12 (Von dem Knesebeck et al., 2005)
Control	1. My age prevents me from doing the things I would like to do	✓	✓	✓
	2. I feel that what happens to me is out of my control	✓	✓	✓
	3. I feel free to plan for the future	✓		
	4. I feel left out of things	✓	✓	✓
Autonomy	5. I can do the things I want to do	✓	✓	✓
	6. Family responsibilities prevent me from doing the things I want to do	✓		✓
	7. I feel that I can please myself what I do	✓	✓	
	8. My health stops me from doing the things I want to do	✓		
	9. Shortage of money stops me from doing things I want to do	✓	✓	✓
Pleasure	10. I look forward to each day	✓	✓	✓
	11. I feel that my life has meaning	✓	✓	✓
	12. I enjoy the things that I do	✓	✓	
	13. I enjoy being in the company of others	✓		
	14. On balance, I look back on my life with a sense of happiness	✓		✓
Self-realization	15. I feel full of energy these days	✓	✓	✓
	16. I choose to do things that I have never done before	✓		
	17. I feel satisfied with the way my life has turned out	✓		
	18. I feel that life is full of opportunities	✓	✓	✓
	19. I feel that the future looks good for me	✓	✓	✓

Note. Items 1,2,4,6,8 and 9 are reverse-coded for the analyses.

3. The empirical evaluation of the CASP scale

Several studies assess the psychometric performance of the CASP-19. The results of the following psychometric properties are reviewed here: 1) internal consistency, 2) factor structure and 3) cross-cultural robustness of the psychometric characteristics.

The internal consistency of each dimension of the CASP-19 scale is examined based on Cronbach's alpha (Hyde et al., 2003; Wiggins et al., 2008; Bowling, 2009; Bowling & Stenner, 2011; Sim et al., 2011). The Cronbach's alpha gives an estimation of the proportion of the total variance that is common among the items of a given dimension. If three indicators reflect the same dimension, they should then share a certain amount of variance. The studies that explore the internal consistency of the CASP-19 scale yield similar results. More specifically, the internal consistency of the pleasure and self-realization dimensions is in general satisfactory, with Cronbach's alpha values ranging from .73 to .83. In contrast, the internal consistency of the control and especially the autonomy dimensions is questionable. For the control dimension, Cronbach's alpha values range from .60 to .64. For the autonomy dimension, Hyde and colleagues (2003) find a Cronbach's alpha of .67, but later studies find values rather close to .50 (Wiggins et al., 2008; Sim et al., 2011).

The factor structure of the CASP-19 scale is tested using confirmatory factor analysis (CFA). The seminal paper of Hyde and colleagues (2003) tests whether the four dimensions are interrelated and measure a common underlying concept of QoL. The authors find strong evidence for a single, underlying QoL factor with strong loadings of the four different dimensions, ranging from .71 to .88. Two later studies compare three factor models: 1) the single-factor model, where all items load on a single latent variable; 2) a first-order factor model, where the items load on their respective dimension and the four dimensions are correlated; and 3) a second-order factor model, where the items load on their respective dimensions and the dimensions in turn load on a second order latent variable, QoL (Wiggins et al., 2008; Sim et al., 2011). In both studies, the results show that the factor structure of the CASP-19 scale does not follow what is predicted by the theory. None of the proposed models fitted the data well (Wiggins et al., 2008; Sim et al., 2011; Vanhoutte, 2012). Namely, there are a number of cross-loadings, which means that some items are not specific to one dimension but are also related to other dimensions, and small loadings, which means that some items only weakly reflect a given dimension.

Based on the prior findings, some authors propose a rectified version of the CASP scale. Wiggins and colleagues (2008) suggest dropping the items that show the weakest correlations with their own dimension and combining the autonomy and the control dimensions. Table 1 displays the items in the CASP-12 version suggested by these authors.

The CASP-12 was tested so far by four studies and the results suggest that it performs better than the CASP-19, but could still be improved (Wiggins et al., Sim et al., 2011, Vanhoutte, 2012; Sexton et al., 2013). Vanhoutte (2012) suggest dropping four items (item 1 "age prevent me from doing", item 6 "family responsibilities", item 8 "my health stops", and item 9 "shortage of money", Table 1) and keeping three dimensions (combining autonomy and control) or two dimensions (combining autonomy, control and self-realization).

Another important aspect in scale validation is the extent to which the measure can be exported to other cultures or countries. Ideally, the psychometric properties of a scale should remain invariant across cultures. Otherwise, cultural differences may simply reflect measurement biases (Chen, 2008). Two studies address this issue for the CASP-19 scale and their results indicate that the psychometric properties of the CASP-19 may not be stable across cultures. Bowling and Stenner (2011) compare the results obtained in the British Omnibus Survey, whose sample is representative of the general British population, with the results obtained in the Ethnibus survey, whose sample is representative of the most common ethnic minorities living in Great Britain. They find that the performance of the CASP-19 is worse in the Ethnibus survey than in the British Omnibus survey. More specifically, a larger number of items fail to meet the criterion for internal consistency in the Ethnibus survey than in the British Omnibus survey. Wu and colleagues (2013) test the psychometric properties of a Chinese version of the CASP-19. Their exploratory factor analysis reveals a fifth dimension and a different distribution of the items across the dimensions. These two studies show that the cross-cultural measurement invariance of the CASP-19 scale cannot be taken as granted. This may also be the case for the CASP-12.

4. The CASP scale used in SHARE

SHARE uses, since 2004, a 12-item version of the CASP scale that is slightly different from the 12-item version suggested later by Wiggins and colleagues (2008). The last column of table 1 shows the items that are included in SHARE. Since this version has not been submitted to empirical validation, the first objective of this paper is to provide a thorough examination of the psychometric properties of the SHARE version of the CASP-12. The second objective of this study is to explore the cross-cultural robustness of the results, because cross-cultural comparison is one of the main purposes of SHARE and can only be performed under measurement invariance (Chen, 2008).

5. Data and methods

5.1 Sample

The analyses are conducted on data from the fourth wave of the Survey of Health, Ageing and Retirement in Europe (SHARE), conducted in 2010/11 (Börsch-Supan, Brandt, Litwin & Weber, 2013; Börsch-Supan, Brandt, Hunkler et al., 2013; Malter & Börsch-Supan, 2013). SHARE is a longitudinal survey conducted every two years since 2004 that collects medical, social and economic data on the population aged 50 and over in Europe. Interviews are conducted face-to-face with the target respondent and willing partners or spouses. The fourth wave is the most recent available, the one with the most participating countries, and includes refreshment samples that considerably increased sample sizes in most countries.

Sixteen countries participated in wave four: Austria (AT), Belgium (BE), Czech Republic (CZ), Denmark (DK), Estonia (EE), France (FR), Germany (DE), Hungary (HU), Italy (IT), Netherlands (NL), Poland (PL), Portugal (PT), Slovenia (SI), Spain (ES), Sweden (SE) and Switzerland (CH) (Table 2). Participants with missing values on one or more of the CASP items are excluded from the analyses. Country-specific sample sizes, sample characteristics, demographic statistics, as well as percentage of individuals who have one or more missing value in the CASP items are reported in table 2. The total sample size amounts to 59'599 individuals (Min=1623; Max=6828) corresponding to 40'685 households (Min=1081; Max=4637). On average, in 52.3 % of the households only one person of the household was interviewed, whereas in 46.3% of the households, two individuals were interviewed. Slightly more than half of the individuals are women (Mean=55.9%, Min=52.9%, Max=59.8%). On average, 27.9% of the respondents live with a partner (Min=21.2%, Max=34.4%). Regarding education, 41.3% (Min=13.2, Max=82.0) have low education, 36.8% (Min=7.1, Max=64.2) have middle education and 21.8% (Min=6.5, Max=41.1) have high education. The average age varies between 64.6 in Hungary and 70.9 in Sweden, with a standard deviation close to 10 years in almost all countries. A small proportion of the sample is below 50 years old, because partners are eligible for the survey regardless of their age. We kept them in the analysis because most of them are close to 50.

Table 2: Country specific sample sizes, sample characteristics and demographic statistics

	Total Sample Size		One interview/ household			Education			Age			One or more missing CASP item
	Individuals N	Households N	%	Female %	Living alone %	Low %	Middle %	High %	Mean	SD	% <50	%
Austria	5332	3795	59.50	57.8	34.40	25.10	49.60	25.30	65.43	10.32	2.7	4.9
Belgium	5388	3778	57.40	55.2	31.20	43.10	26.60	30.30	65.18	11.16	2.4	5.5
Czechia	6196	4185	52.00	57.7	29.90	45.70	42.10	12.20	65.29	9.96	2.2	5.8
Denmark	2393	1616	52.00	54.5	28.60	19.60	39.30	41.10	65.47	11.36	2.1	3.6
Estonia	6828	4637	52.70	59.8	31.00	26.00	35.70	38.30	66.65	10.28	1.6	10.0
France	5954	4205	58.40	56.9	32.00	45.50	34.10	20.40	66.01	11.37	3.1	8.1
Germany	1623	1081	50.20	52.9	23.70	13.20	56.20	30.50	68.44	9.02	.6	4.8
Hungary	3076	2021	47.80	57.0	27.50	31.30	52.50	16.20	64.65	9.80	2.4	2.3
Italy	3673	2333	43.00	54.8	21.70	70.60	22.90	6.50	66.97	10.06	1.7	3.6
Netherlands	2822	1912	52.50	55.6	23.90	47.50	26.20	26.30	66.18	9.96	1.4	4.7
Poland	1880	1161	38.20	55.3	31.80	44.70	47.60	7.70	67.25	9.70	.6	3.9
Portugal	2080	1384	49.70	57.0	21.20	64.40	7.10	28.50	64.71	9.96	2.8	4.2
Slovenia	2756	2117	69.80	56.6	25.80	34.50	49.10	16.30	65.28	10.21	1.5	3.9
Spain	3690	2316	42.30	54.9	24.50	82.00	9.30	8.70	68.23	11.42	1.7	6.0
Sweden	2122	1490	57.60	53.8	33.80	48.20	27.00	24.90	70.89	10.04	.3	3.7
Switzerland	3786	2654	57.40	55.0	24.70	20.00	64.20	15.80	65.08	10.64	2.7	1.8
Total	**59599**	**40685**										
Average	**3725**	**2543**	**52.5**	**55.9**	**27.9**	**41.3**	**36.8**	**21.8**	**66.4**	**10.3**	**1.9**	**4.8**
Minimum	**1623**	**1081**	**38.2**	**52.9**	**21.2**	**13.2**	**7.1**	**6.5**	**64.6**	**9.0**	**0.3**	**1.8**
Maximum	**6828**	**4637**	**69.8**	**59.8**	**34.4**	**82.0**	**64.2**	**41.1**	**70.9**	**11.4**	**3.1**	**10.0**

5.2 The SHARE version of the CASP-12

The items included in the SHARE version of the CASP-12 are listed in table 1. Answers are coded on a 4-point Likert scale (1. Often, 2. Sometimes, 3. Rarely and 4. Never). All items are (re)coded in such a way that higher scores indicate a higher level of QoL.

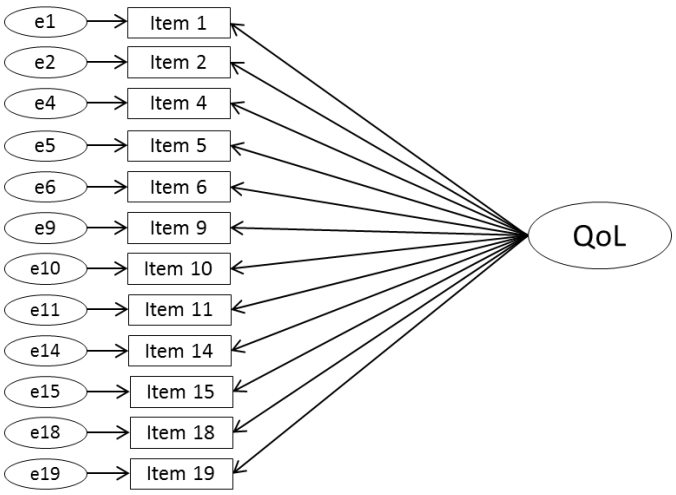
5.3 Statistical analyses

All analyses are conducted separately by country. As a first-step, internal consistency is assessed for each dimension using the ordinal Cronbach's alpha coefficient4 (Gadermann et al., 2012; Zumbo et al., 2007). The ordinal Cronbach's alpha can take any value between 0 and 1. Higher values indicate that scores on the considered dimension are internally consistent. Following DeVellis (2003), internal consistency is considered unacceptable if the alpha value is below .60, undesirable if between .60 and .65, minimally acceptable if between .65 and .70 and respectable between .70 and .80.

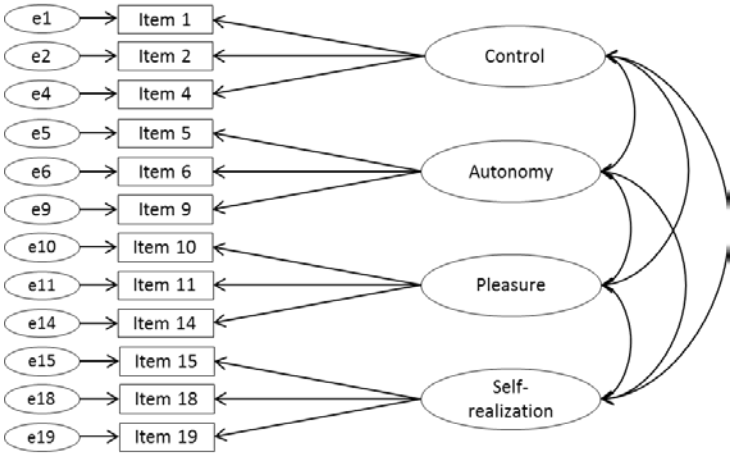
In a second step, first- and second-order CFA are conducted, to examine the factor structure of the SHARE version of the CASP-12. The following three factor models, originally presented in Wiggins and colleagues (2008), are tested (Figure 1):

1. Model A: single-factor model where all indicators load on a single latent variable representing overall QoL.
2. Model B: four-dimension first-order factor model where the indicators load on their own dimension and the four dimensions are allowed to correlate with one another.
3. Model C: second-order factor model where the indicators load on their own dimension and the four dimensions in turn load on a higher-order latent variable representing overall QoL.

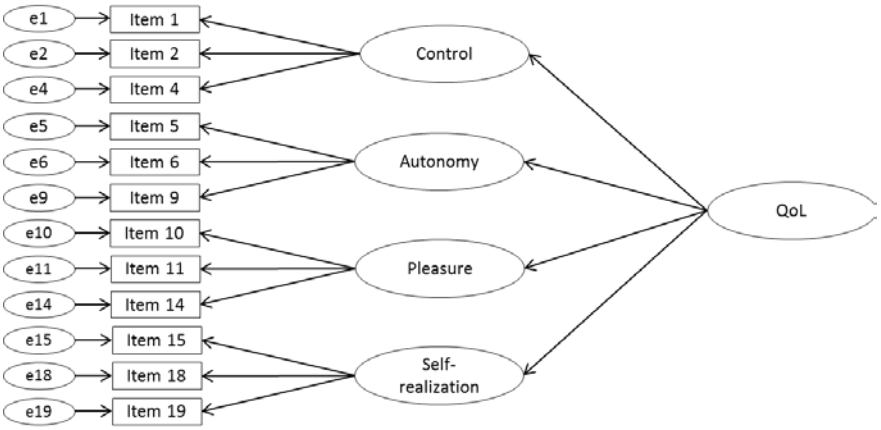
[4] The computation of Cronbach's alpha is usually based on a Pearson correlation matrix. However, an important assumption for the use of Pearson correlations is that variables are continuous. For binomial or ordinal indicators, Zumbo et al (2007) developed an ordinal Cronbach's alpha based on tetrachoric or polychoric correlations.



a) Model A: the single factor model



b) Model B: the four dimensions first-order factor model



c) Model C: the four dimensions second-order factor model

Figure 1: The different models as originally, adapted from Wiggins and colleagues (2008)

Version 5.1 of Mplus is used for the factor analyses. The estimator used is the Weighted Least Squares Means and Variances adjusted estimator (WLSMV), because it does not assume normally distributed variables and provides the best option for modelling categorical data (Brown, 2006).

We use three types of global fit measures as criteria to decide which model fits the data better: the Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI), and the Trucker-Lewis Index (TLI). The RMSEA reflects the degree to which a model fits the population covariance matrix while taking into account the degrees of freedom and sample size (Brown, 2006). It is a parsimony-adjusted index that favors simpler models. Values lower than .08 indicate a good model fit (Hu & Bentler, 1999; Browne & Cudeck, 1993). The CFI and the TLI compare the fit of the model to a more restricted, baseline model. Like the RMSEA, the TLI favors simpler models, which is not the case for the CFI. CFI and TLI values greater than .95 indicate a good model fit (Hu & Bentler, 1999)5. We consider that the model fit is good if all fit indices are good according to the standard cut-off criteria, marginal if one of the fit indices is bad, and poor if two or all three fit indices are bad. In addition to the fit indices, standardized factor loadings, residual variances and modification indices are examined to identify the sources of ill fit.

Spearman correlations between each item and the total score of each dimension are computed. If the factor structure postulated by the theory is correct, then each item should correlate in the expected direction with all dimensions and have a stronger correlation with its own dimension than with the other dimensions.

Finally, to propose a revised CASP scale using the items available in SHARE, we conduct exploratory factor analyses (EFA). Oblique rotation is applied to obtain a meaningful interpretation of the factors. We rely on Kaiser's criterion and scree plots to determine the optimal number of factors (add reference).

6. Results

The tables presented in this paper report a summary of the results that were obtained across all 16 countries. Results for each country are available upon request from the first authors.

[5] The chi-square statistics are also commonly used in the literature. However, they are not relied upon in this study, because they are inflated in large samples, like the ones we have in SHARE, resulting in a higher probability to reject the model. The Standardized Root Mean Square Residual (SRMR) is also very common but is not available for models with categorical indicators (Yu, 2002).

6.1 Reliability analysis

The ordinal Cronbach's alpha values listed in table 3 shows that internal consistency varies from one dimension to another.

Table 3: Cronbach's alpha for each dimension by country

	Control	Autonomy	Pleasure	Self-real.	Control & Autonomy
Austria	0.81	0.33	0.85	0.88	0.73
Belgium	0.70	0.39	0.67	0.82	0.68
Czechia	0.79	0.29	0.77	0.82	0.68
Denmark	0.67	0.36	0.84	0.88	0.66
Estonia	0.78	0.29	0.75	0.85	0.71
France	0.74	0.28	0.80	0.83	0.66
Germany	0.53	0.54	0.82	0.79	0.66
Hungary	0.72	0.22	0.80	0.89	0.65
Italy	0.85	0.48	0.34	0.86	0.76
Netherlands	0.62	0.29	0.76	0.85	0.65
Poland	0.80	0.22	0.89	0.92	0.69
Portugal	0.69	0.25	0.38	0.83	0.65
Slovenia	0.78	0.34	0.82	0.88	0.67
Spain	0.79	0.27	0.79	0.85	0.68
Sweden	0.67	0.32	0.74	0.85	0.66
Switzerland	0.69	0.34	0.81	0.81	0.67
Average	**0.73**	**0.33**	**0.74**	**0.85**	**0.68**
SD	**0.08**	**0.09**	**0.16**	**0.03**	**0.03**
Min	**0.53**	**0.22**	**0.34**	**0.79**	**0.65**
Max	**0.85**	**0.54**	**0.89**	**0.92**	**0.76**

The reliability coefficient is unacceptably weak for the **autonomy dimension** (average α=.33, *SD*=.09, *Min*=.22, *Max*=.54) and satisfactory for the **self-realization dimension** (average α =.85, *SD*=.03, *Min*=.79, *Max*=.92). The internal consistency of the **control dimension** is on average satisfactory (average α =.73, *SD*=.08, *Min*=.53, *Max*=.85), and in most countries it is above or very close to the cut-off of .70. Only Germany scores lower (α =.53), which reveals weak internal consistency of the control dimension in this country. An examination of the correlation matrix shows that this is principally due to item 4 ("I feel left out of things"), which correlates weakly with item 1 ("My age prevents me from doing"; *r*=.191), and item 2 ("If feel that what happens is out of my control"; *r*=.155). With respect to the **pleasure dimension**, the Cronbach's alpha is on average satisfactory (average α =.74, *SD*=.16, *Min*=.34, *Max*=.89) and above .70 in most countries. However, it is very low in Italy and Portugal (respectively

α =.34 and α=.38), revealing weak internal consistency of the pleasure dimension in these two countries. In both countries, this is due to item 10 ("I look forward to each day"), which correlates weakly and sometimes negatively with item 11 ("I feel that my life has meaning"; r_{Italy}= -.134; $r_{Portugal}$= -.039), and item 14 ("I look back on my life with a sense of happiness"; r_{Italy}= -.172; $r_{Portugal}$=.030).

6.2 First- and second-order CFA of the original models

Table 4 shows the CFA results for the single-factor model (Model A), the first-order model (Model B) and the second-order model (Model C). In general, model fit is poor. First, Model A does not fit the data well in any country. Model B only fits the data well in Austria (CFI=.955, TLI=.977, RMSEA=.079) and Slovenia (CFI=.962, TLI=.977, RMSEA=.072), and marginally well in Switzerland (CFI=.913, TLI=.953, RMSEA=.073) and Poland (CFI=.956, TLI=.977, RMSEA=.104). In all other countries, at least two of the three fit indices considered point toward a bad fit. The results for Model C are very similar to the results for Model B, with a bad fit in most countries, a good fit in Austria (CFI=.950, TLI=.976, RMSEA=.079) and Slovenia (CFI=.955, TLI=.974, RMSEA=.077), and a marginally good fit in Poland (CFI=.951, TLI=.976, RMSEA=.106).

The most important source of ill fit seems to be related to the autonomy dimension. First, in eight countries (AT, CZ, ES, FR, HU, NL, PL, PT), the latent variable correlation matrix is non-positive definite, which is due to high multicollinearity between the autonomy dimension and the other dimensions of the model B and C. Second, the standardized loadings of the autonomy dimension, and especially those of items 6 ("family responsibilities") and 9 ("shortage of money"), are relatively weak in comparison to the loadings of the other dimensions (Table 5). This is also the case for Austria and Slovenia, the two countries where fit is generally good. Finally, according to the modification indices, model fit could be significantly improved by adding cross-loadings from the autonomy indicators onto other dimensions. This was especially the case for item 5 ("I can do the things I want to do").

Table 4: Fit indices of the first-order models, by country*

	Model A			Model B			Model C		
	CFI	TLI	RMSEA	CFI	TLI	RMSEA	CFI	TLI	RMSEA
Austria	0.873	0.939	0.126	0.955	0.977	0.076	0.950	0.976	0.079
Belgium	0.782	0.857	0.128	0.90	0.929	0.090	0.882	0.918	0.097
Czechia	0.726	0.828	0.181	0.933	0.958	0.090	0.930	0.957	0.091
Denmark	0.775	0.879	0.147	0.848	0.811	0.126	0.844	0.911	0.126
Estonia	0.82	0.894	0.141	0.941	0.964	0.082	0.921	0.952	0.094
France	0.831	0.907	0.126	0.922	0.954	0.089	0.916	0.953	0.090
Germany	0.835	0.89	0.114	0.917	0.942	0.083	0.890	0.923	0.096
Hungary	0.846	0.909	0.148	0.932	0.957	0.101	0.929	0.957	0.102
Italy	0.792	0.867	0.199	0.892	0.933	0.141	0.886	0.929	0.145
Netherlands	0.832	0.892	0.106	0.909	0.935	0.082	0.891	0.926	0.087
Poland	0.867	0.931	0.179	0.956	0.977	0.104	0.951	0.976	0.106
Portugal	0.666	0.726	0.199	0.739	0.786	0.176	0.737	0.792	0.174
Slovenia	0.878	0.933	0.123	0.962	0.977	0.072	0.955	0.974	0.077
Spain	0.856	0.924	0.143	0.944	0.969	0.092	0.942	0.969	0.092
Sweden	0.823	0.894	0.124	0.905	0.939	0.094	0.897	0.935	0.097
Switzerland	0.824	0.910	0.101	0.913	0.953	0.073	0.894	0.944	0.079

* To facilitate reading, satisfactory values are highlighted in green

Table 5: Descriptive statistics on the country specific standardized factor loadings

Model A	CASP											
	item1	item2	item4	item5	item6	item9	item10	item11	item14	item15	item18	item19
Average	.53	.58	.56	.50	.17	.28	.57	.77	.57	.77	.79	.80
SD	(.06)	(.10)	(.13)	(.10)	(.06)	(.08)	(.34)	(.06)	(.11)	(.05)	(.06)	(.04)
25% perc.	.50	.50	.48	.45	.14	.21	.59	.73	.49	.74	.75	.78
50% perc.	.52	.55	.57	.51	.15	.29	.71	.78	.57	.76	.79	.80
Minimum	.42	.45	.34	.31	.08	.14	-.31	.62	.40	.70	.69	.73
Maximum	.62	.74	.75	.66	.31	.42	.82	.85	.80	.86	.90	.85

Model B	Control			Autonomy			Pleasure			Self-realization		
	item1	item2	item4	item5	item6	item9	item10	item11	item14	item15	item18	item19
Average	.67	.74	.70	.52	.19	.31	.64	.86	.65	.80	.82	.83
SD	(.08)	(.08)	(.13)	(.15)	(.10)	(.11)	(.36)	(.04)	(.11)	(.04)	(.05)	(.03)
25% perc.	.61	.67	.64	.41	.13	.21	.68	.84	.59	.77	.78	.81
50% perc.	.67	.72	.71	.50	.15	.27	.79	.86	.64	.79	.83	.83
Minimum	.55	.61	.46	.33	.10	.14	-.33	.80	.46	.73	.74	.78
Maximum	.78	.88	.86	.94	.43	.51	.88	.92	.89	.89	.92	.87

Model C	CASP											
	Control			Autonomy			Pleasure			Self-realization		
Average	.67			1.04			.84			.96		
SD	(.07)			(.29)			(.04)			(.03)		
25% perc.	.64			.87			.80			.94		
50% perc.	.67			1.00			.84			.97		
Minimum	.55			.59			.77			.55		
Maximum	.82			1.68			.91			1.02		
	item1	item2	item4	item5	item6	item9	item10	item11	item14	item15	item18	item19
Average	.67	.73	.70	.53	.18	.30	.63	.87	.65	.80	.82	.83
SD	(.08)	(.08)	(.13)	(.16)	(.09)	(.11)	(.36)	(.04)	(.11)	(.04)	(.05)	(.03)
25% perc.	.61	.67	.64	.42	.12	.20	.68	.84	.59	.77	.78	.81
50% perc.	.67	.72	.71	.52	.14	.27	.79	.86	.64	.79	.83	.83
Minimum	.55	.61	.48	.33	.09	.14	-.33	.80	.46	.73	.74	.78
Maximum	.78	.88	.86	.94	.41	.50	.88	.92	.89	.89	.92	.87

6.3 Item-total Spearman correlations

Spearman item-total correlations allow to explore further the factor structure of the CASP-12 and to better understand the sources of ill-fit. As a reminder, if the factor structure postulated by the theory is correct, then each item should correlate in the expected direction with all dimensions and have a stronger correlation with its own dimension than with the other dimensions. Table 6 displays the summary statistics of the 16 country-specific correlation matrices. **For the control dimension**, all items have a moderate to strong correlation with their own dimension, and the items do not correlate more strongly with other dimensions, in most cases. There are only two exceptions: in Sweden and Spain, item 1 ("My age prevents me from") correlates slightly more strongly with the self-realization dimension than with the control dimension. **For the autonomy dimension**, the results are more problematic. First, all three items have low correlations with their own dimension (item 5: *M*=.077, *SD*=.045; item 6: *M*=.158, *SD*=.060; item 9: *M*=.211, *SD*=.057). Moreover, item 5 correlates moderately with the other dimensions, and in all countries it has a stronger correlation with the self-realization dimension. This suggests that item 5 ("I can do the things I want to do") should rather be assigned to the self-realization dimension. In contrast, item 6 ("Family responsibilities prevent me") and item 9 ("Shortage of money stops me from doing things I want to do") correlate only weakly with the other dimensions, and in most countries the strongest —although weak— correlation is found with the control dimension. This result suggests that the two items, in general, do not share any common variance with the other items. **With respect to the pleasure dimension**, all items have on average a moderate to strong correlation with their own dimension (Item 10: *M*=.416, *SD*=.216; Item 11: *M*=.490, *SD*=.134; Item 14: *M*=.761, *SD*=.064). In some countries, some items correlate more strongly with the self-realization dimension than with the pleasure dimension. This happens in one country for item 10 (FR) and in six countries for item 11 (BE, IT, NL, PT, SI, SE). It is also worth noticing that two countries, Italy and Portugal, stand out from the correlation pattern that is found for the other countries. In these two countries, item 10 ("I look forward to each day") correlates weakly and negatively with the pleasure dimension (IT: r_{item10}=-.161; PT: r_{item10}=-.041) and negatively with the other dimensions. Moreover, in comparison to the other countries, item 11 ("I feel that my life has meaning") has a very low correlation with the pleasure dimension (IT: r_{item11}=.238; PT: r_{item11}=.201). Finally, for **the self-realization dimension**, all items have on average a strong correlation with their own dimension (Item 15: *M*=.593, *SD*=.067; Item 18: *M*=.654, *SD*=.073; Item 19: *M*=.838, *SD*=.027) and the correlations with the other dimensions are weaker.

Table 6: Summary statistics of the item-total Spearman correlations

Items	Control				Autonomy				Pleasure				Self-Realization			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
item 1	**.435**	**.067**	**.328**	**.565**	.236	.058	.135	.326	.189	.075	.022	.275	.375	.053	.294	.458
item 2	**.542**	**.080**	**.416**	**.686**	.299	.041	.186	.358	.216	.102	-.040	.359	.339	.071	.251	.475
item 4	**.466**	**.094**	**.313**	**.638**	.293	.052	.220	.388	.254	.105	-.009	.450	.317	.094	.152	.472
item 5	.268	.055	.173	.344	**.077**	**.045**	**-.019**	**.168**	.285	.055	.183	.401	.386	.063	.237	.449
item 6	.210	.044	.152	.309	**.158**	**.060**	**.089**	**.332**	.051	.062	-.103	.138	.039	.045	-.032	.128
item 9	.234	.057	.120	.308	**.211**	**.057**	**.117**	**.341**	.122	.089	-.084	.226	.180	.064	.061	.260
item 10	.185	.201	-.303	.390	.138	.166	-.281	.266	**.416**	**.216**	**-.161**	**.659**	.376	.207	-.184	.574
item 11	.303	.066	.189	.427	.212	.052	.127	.302	**.490**	**.134**	**.201**	**.711**	.507	.057	.407	.616
item 14	.207	.075	.087	.361	.198	.061	.102	.292	**.394**	**.102**	**.199**	**.664**	.424	.094	.319	.627
item 15	.419	.060	.344	.538	.236	.043	.156	.307	.434	.076	.314	.571	**.593**	**.067**	**.511**	**.755**
item 18	.344	.067	.221	.480	.233	.041	.147	.315	.468	.080	.289	.619	**.654**	**.073**	**.505**	**.799**
item 19	.363	.059	.259	.465	.265	.047	.165	.326	.487	.081	.302	.606	**.636**	**.053**	**.531**	**.725**

6.4 A 10-item version

The previous results suggest that item 6 ("family responsibilities") and 9 ("shortage of money") are bad indicators of the SHARE version of the CASP-12. Therefore, we drop these two items and examine the factor structure of the 10 remaining items, using EFA and CFA.

The results of the EFA suggest a two-factor structure in all countries. Systematically, two factors have eigenvalues over Kaiser's criterion of 1 and the scree plots show a clear inflexion point after the second factor. The distribution of the items between the two factors is also very consistent. As shown in table 7, items 1, 2, and 4 systematically load on the first factor, whereas items 5, 10, 11, 14, 15, 18, and 19 load on the second factor. Thus, the Control dimension is kept intact and all the other items are merged into a single dimension that can be labelled "Positivity".

Table 7: Factor loadings of the exploratory factor analysis with 10 items

	First factor				Second factor			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Item 1	0.545	0.09	0.405	0.681	0.146	0.08	0.020	0.312
Item 2	0.824	0.06	0.722	0.919	-0.013	0.02	-0.054	0.009
Item 4	0.613	0.12	0.435	0.831	0.109	0.09	-0.020	0.314
Item 5	0.124	0.07	-0.032	0.234	0.428	0.09	0.226	0.568
Item 10	-0.102	0.14	-0.433	0.050	0.661	0.27	-0.001	0.879
Item 11	-0.025	0.07	-0.136	0.147	0.810	0.06	0.675	0.879
Item 14	-0.073	0.05	-0.167	0.010	0.635	0.09	0.509	0.884
Item 15	0.211	0.07	0.069	0.305	0.649	0.06	0.555	0.766
Item 18	0.050	0.08	-0.147	0.166	0.777	0.07	0.667	0.948
Item 19	0.067	0.08	-0.106	0.196	0.774	0.06	0.648	0.898

We test the two-factor structure suggested by the exploratory analyses with CFA. Table 8 displays the general fit indices for the CFA. The results show that, in general, the first-order two-factor model fits the data better than the single factor model. A second-order factor model cannot be tested because at least three first-order factors (dimensions) would be required to fit such a model. Table 9 shows the standardized factor loadings for each country. They are all significant and almost all of them are above .40. In some countries, the fit is marginally worse due to cross-loadings of one or more items. Portugal and Italy remain two exceptions due to the problems related to item 10, already identified in the previous analyses. The correlation between the two factors is on average .63 (*SD*=.07, *Min*=.477, *Max*=.781).

Table 8: Fit indices for the confirmatory factor analysis based on 10 items

	Single factor model			First-order factor model		
	CFI	TLI	RMSEA	CFI	TLI	RMSEA
Austria	0.874	0.942	0.144	0.960	0.983	0.078
Belgium	0.818	0.879	0.136	0.941	0.961	0.077
Czechia	0.717	0.811	0.224	0.904	0.944	0.122
Denmark	0.886	0.938	0.124	0.969	0.983	0.064
Estonia	0.825	0.899	0.163	0.934	0.963	0.098
France	0.846	0.920	0.139	0.955	0.977	0.075
Germany	0.843	0.907	0.123	0.861	0.915	0.118
Hungary	0.847	0.914	0.171	0.923	0.957	0.121
Italy	0.832	0.888	0.214	0.941	0.97	0.110
Netherlands	0.884	0.930	0.101	0.945	0.967	0.070
Poland	0.864	0.932	0.213	0.927	0.968	0.149
Portugal	0.734	0.794	0.200	0.847	0.891	0.145
Slovenia	0.880	0.937	0.143	0.961	0.980	0.080
Spain	0.860	0.924	0.172	0.927	0.962	0.122
Sweden	0.844	0.913	0.134	0.923	0.957	0.094
Switzerland	0.860	0.930	0.102	0.958	0.980	0.055

* To facilitate reading, satisfactory values are highlighted in green

Table 9: Factor loadings of the first-order factor model with 10 items

		First factor			Second factor							Correlation between factors
		Item 1	Item 2	Item 4	Item 5	Item 10	Item 11	Item 14	Item 15	Item 18	Item 19	
AT	B	.68	.78	.86	.53	.78	.84	.67	.82	.82	.85	.71
	SE	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
BE	B	.56	.70	.75	.55	.46	.71	.50	.72	.78	.82	.57
	SE	(.02)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.02)
CZ	B	.66	.88	.80	.43	.76	.81	.44	.76	.73	.77	.53
	SE	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
DK	B	.56	.70	.67	.61	.78	.83	.64	.79	.85	.86	.59
	SE	(.02)	(.02)	(.02)	(.02)	(.02)	(.02)	(.02)	(.01)	(.01)	(.01)	(.02)
EE	B	.74	.66	.82	.52	.63	.74	.56	.78	.83	.79	.63
	SE	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
FR	B	.64	.73	.73	.56	.79	.80	.56	.77	.78	.81	.63
	SE	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
DE	B	.49	.62	.57	.64	.71	.76	.63	.72	.72	.75	.78
	SE	(.03)	(.03)	(.03)	(.02)	(.03)	(.02)	(.02)	(.02)	(.02)	(.02)	(.03)
HU	B	.79	.67	.61	.42	.65	.80	.60	.82	.87	.85	.59
	SE	(.02)	(.02)	(.02)	(.02)	(.01)	(.01)	(.02)	(.01)	(.01)	(.01)	(.02)
IT	B	.72	.85	.86	.50	-.29	.78	.83	.82	.81	.81	.62
	SE	(.01)	(.01)	(.01)	(.02)	(.02)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
NL	B	.62	.66	.51	.61	.58	.75	.56	.78	.84	.81	.62
	SE	(.02)	(.02)	(.02)	(.02)	(.02)	(.02)	(.02)	(.01)	(.01)	(.01)	(.02)
PL	B	.63	.82	.84	.52	.77	.85	.79	.87	.90	.84	.68
	SE	(.02)	(.02)	(.02)	(.02)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.02)
PT	B	.80	.70	.47	.58	-.14	.66	.51	.78	.79	.82	.48
	SE	(.02)	(.02)	(.03)	(.02)	(.02)	(.02)	(.02)	(.01)	(.01)	(.01)	(.03)
SI	B	.75	.85	.64	.31	.77	.84	.64	.82	.87	.84	.67
	SE	(.02)	(.01)	(.02)	(.02)	(.01)	(.01)	(.02)	(.01)	(.01)	(.01)	(.02)
ES	B	.71	.80	.76	.52	.83	.87	.43	.82	.74	.80	.70
	SE	(.01)	(.01)	(.01)	(.02)	(.01)	(.01)	(.02)	(.01)	(.01)	(.01)	(.01)
SE	B	.69	.60	.64	.65	.65	.74	.47	.76	.84	.83	.62
	SE	(.02)	(.02)	(.02)	(.02)	(.02)	(.02)	(.02)	(.01)	(.01)	(.01)	(.02)
CH	B	.63	.72	.66	.38	.80	.81	.61	.73	.75	.78	.66
	SE	(.02)	(.02)	(.02)	(.02)	(.01)	(.01)	(.02)	(.01)	(.01)	(.01)	(.02)

7. Discussion and conclusion

The CASP scale was developed to measure the quality of life (QoL) of older individuals. The purpose of this article is to examine the psychometric properties of the SHARE version of the CASP-12 and to investigate its cross-cultural robustness. Analyses are based on the fourth wave of SHARE, which includes 16 countries.

Following Wiggins et al. (2008), three models are tested using CFA: 1) a single factor model, 2) a four-dimension first-order model, and 3) a second-order factor model. In the second-order factor model, a single second-order latent variable reflects the different CASP dimensions and can be interpreted as QoL. This means that in practice, the CASP items can be combined into a measure of overall QoL, which is what is assumed by the theoretical approach used by the authors to develop their CASP scale (Higgs et al., 2003).

The results show that none of the three theoretical models fits the data well. Though there are differences across countries, there is a general pattern. The most important source of ill fit is related to the autonomy dimension, which shows very low internal consistency. Moreover, two of its indicators, items 6 ("family responsibilities") and 9 ("shortage of money"), seem to be bad indicators of the SHARE version of the CASP-12. They have much weaker factor loadings than the other CASP items and do not correlate with any dimension, not even their own. Finally, item 5 correlates more strongly with the self-realization dimension than with the autonomy dimension.

Based on the aforementioned results, we drop the two problematic items from the scale and explore the factor structure of the ten remaining items using EFA and CFA. The EFA suggests a two-factor structure in all countries. However whereas a first-order two-factor model fits the data better than a single factor model, it doesn't perform extremely well in all countries due to cross-loadings of one or more items. These results suggest that additional country specific modifications are needed.

Several reasons may explain why the results do not conform to the theoretical expectations. First, this may be due to some specific characteristics of the samples. The CASP scale has been developed based on a small sample of 286 British individuals aged 65 to 75. Although the authors state that their sample was representative of those of the same age in the British population, their sample may nevertheless have been very different from the population of a representative large-scale survey. Second, the set of items used in SHARE may explain some of the discrepant results. As a reminder, SHARE uses a different set and a reduced number of items and this may have impaired our ability to model correctly the different dimensions separately. And third, the theoretical background of the CASP scale may

present some weaknesses. This may be the case for the autonomy dimension given that two other studies found similar findings with respect to items 6 ("family responsibilities") and 9 ("shortage of money") (Sim et al., 2011; Vanhoutte, 2012). As a reminder, factor analyses are based on correlation patterns, and to work well, an item has to be moderately or strongly correlated with the other items of the same dimension. In the case of the autonomy dimension, this means that the authors of the CASP scale expected that people who are annoyed by family responsibilities are also partially annoyed by financial issues. In our view, this expected correlation is highly questionable. Moreover, it is worth noticing that these two items are very specific compared to the other items of the scale. They specifically relate the lack of autonomy to financial and family issues; whereas the other items remain very general, leaving respondents the freedom to relate the lack of autonomy, control, self-realization and pleasure to whatever is relevant for them. Thus, these two items may discriminate a very specific group of individuals (i.e. those who have financial problems) which is different from those individuals which will be discriminated by the other –more general– items of the scale.

The two factors solution we found with the ten remaining items is not comparable with the two-factor solution suggested by Vanhoutte (2012), because the analytical approach is very different. In this paper, we test the models suggested by the authors of the CASP scale, dropped the problematic items and explored the factor structure of the 10 remaining items statistically. Vanhoutte (2012) defines ex-ante his two-factor solution based on theoretical reflections. In his two factor solution, the pleasure dimension forms a separate dimension, called hedonic dimension, and all other items load on a unique dimension, called eudemonic dimension. However, our results, and more specifically the item-total correlation matrix, rather suggest that the pleasure dimension is very close to the self-realization dimension. The items of both dimensions load in all countries on the same factor and the items of the control dimension load clearly on a separate dimension. Moreover, Vanhoutte (2012) also suggest to drop item 1 ("My age prevents me from doing the things I would like to do") because it refers to age which, according to him, is a driver of quality of life and should not be included in the measure of the outcome. We decided not to drop this item, because there was no reason to do so at the statistical level, and at theoretical level, age per se is not considered as a strong predictor of QoL (Blane et al., 2008).A closer look at country specificities shows that Italy and Portugal stand out of the general results pattern with respect to the pleasure dimension. More specifically, item 10 ("I look forward to each day") correlates negatively with all dimensions. This is related to the translation of this item in these countries. Item 10 "I look forward to each

day" has an extremely positive connotation. However, the expression "I look forward" is difficult to translate in some languages. In Italy and Portugal, it seems that, without any contextual information, the wording that is used was sometimes interpreted positively and sometimes negatively. This result underlines the importance of translation and shows how small imprecisions may lead to misunderstanding and consequently to lack of internal consistency.

The analyses we present in this paper are usually the first step to check the invariance of a scale. However, it presents a strong methodological limitation because the 10-items version we suggested was tested on the same sample as the one we used to examine the validity of the 12-items version. This methodological approach is questionable. However, the fact that we ran the same analyses across 16 countries and that the general result pattern (namely the fact that the financial and family items do not work well) is the same suggest that our conclusion is not completely mistaken. Nevertheless, our 10-items version still needs to be validated on a different sample. Future research can also follow up by examining further 1) the configural invariance, 2) the metric invariance and 3) the scalar invariance, using Multiple Group Confirmatory Factor Analysis. Given the weaknesses of the model initially postulated by the theory, any further check of measurement invariance should drop from the model item 6 and 9. Future research should also take into consideration the fact that respondents are nested in couples. Finally, it would be interesting to examine more in detail the applicability of the scale to old-old individuals. The CASP scale was developed for retirees aged 65 to 75, who are still healthy enough to bite into life. However, the ageing process is very heterogeneous and there is no guarantee that the same scale applies to very old and unhealthy individuals. Thus, future research should examine this aspect more in detail, comparing the results obtained among young retirees with the results obtained among old-old individuals.

References

Blane, D., Netuveli, G., & Montgomery, S.M. (2008). Quality of life, health and physiological status and change at older ages. *Social Science & Medicine*, *66*(7), 1579-87.

Bowling, A. (2009). The psychometric properties of the older people's quality of life questionnaire, compared with the CASP-19 and the WHOQOL-OLD. *Current Gerontology and Geriatrics Research*, 1-12.

Bowling, A., & Stenner, P. (2011). Which measure of quality of life performs best in older age? A comparison of the OPQOL, CASP-19 and WHOQOL-OLD. *Journal of Epidemiology Community Health*, *65*, 273-280.

Brown, T.A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York, London: The Guilford Press.

Brown, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds.), *Testing Structural Equation Models* (pp. 136–162). Beverly Hills, CA: Sage

Börsch-Supan A., M. Brandt, H. Litwin and G. Weber (Eds). (2013). *Active ageing and solidarity between generations in Europe: First results from SHARE after the economic crisis.* Berlin: De Gruyter.

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., Schaan, B., Stuck, S., & Zuber, S. (2013). Data resource profile: The survey of health, ageing and retirement in Europe (SHARE). *International Journal of Epidemiology*, *42*(4), 992-1001.

Chen, F.F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Attitudes and Social Cognition*, *95*(5), 1005-1018.

Deci, E. L., & Ryan, R. M. (1987). The support of autonomy and the control of behavior. *Journal of Personality and Social Psychology*, *57*(6), 1024-1037.

DeVellis, R. F. (2003). *Scale Development. Theory and Applications*. Thousand Oaks, London, New Delhi: Sage.

Doyal, L., & Gough, I. (1991). *A theory of human need*. Hong Kong: Macmillan.

Flora D.B., & Curran, P.J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, *9*(4), 466-491.

Gardemann, A.M., Guhn, M., & Zumbo, B.D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation, 17*(3), 1-13.

Higgs, P., Hyde, M., Wiggins, R., & Blane, D. (2003). Researching quality of life in early old age: The importance of the sociological dimension. *Social Policy and Administration, 37*(3), 239-252.

Howel, D. (2012). Interpreting and evaluating the CASP-19 quality of life measure in older people. *Age and Ageing*, 0, 1-6.

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.

Hutchinson, S.R., & Olmos, A. (1998). Behavior of descriptive fit indexes in confirmatory factor analysis using ordered categorical data. *Structural Equation Modeling: A Multidisciplinary Journal, 5*, 344-364.

Hyde, M., Wiggins, R., Higgs, P., & Blane, D. (2003). A measure of quality of life in early old age: The theory, development and properties of needs satisfaction model. *Aging and Mental Health, 7*(3), 186-194.

Malter, F., Börsch-Supan, A.(Eds.) (2013). *SHARE Wave 4: Innovations & Methodology*. Munich: MEA, Max Planck Institute for Social Law and Social Policy.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review, 50*, 370-396.

Rotter, J. B., & Mulry, R. C. (1965). Internal versus external control of reinforcement and decision time. *Journal of Personality and Social Psychology, 2*(4), 598-604.

Sexton, E., King-Kallimanis, B.L., Conroy, R.M., & Hickey, A. (2013). Psychometric evaluation of the CASP-19 quality of life scale in an older Irish cohort. *Quality of Life Research, 22*(9), 2549-2559.

Sim, J., Bartlam, B., & Bernard, M. (2011). The CASP-19 as a measure of quality of life in old age: evaluation of its use in a retirement community. *Quality of Life Research*, *20*(7), 997-1004.

Vanhoutte, B. (2012). *Measuring subjective well-being in later life: A review*. Manchester.

von dem Knesebeck, O., Hyde, M., Higgs, P., Kupfer, A., & Siegriest, J. (2005). Quality of life and well-being In A. Börsch-Supan, A. Brugiavini, H. Jürges, J. Mackenbach, J. Siegriest & B. Weaver (Eds.), *Health, ageing and retirement in Europe: First results from the survey of health, ageing and retirement in Europe* (pp. 199-2003). Mannheim, Germany: Strauss.

Wiggins, R., Netuveli, G., Hyde, M., Higgs, P., & Blane, D. (2008). The evaluation of a self-enumerated scale of quality of life (CASP 19) in the context of research on Ageing: A combination of explanatory and confirmatory approaches. *Social Indicators Research*, *89*, 61-77.

Wu, T.-Y., Chie, W.-C., Kuo, K.-L., Wongb, W.-K., Liu, J.-P., Chiu, S.-T., et al. (2013). Quality of life (QOL) among community dwelling older people in Taiwan measured by the CASP-19, an index to capture QOL in old age. *Archives of Gerontology and Geriatrics, 57*, 143-150.

Yu, C.-Y. (2002). Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes. Unpublished doctoral dissertation. Retrieved on January 2014, from http://www.statmodel.com/download/Yudissertation.pdf

Zumbo, B.D., Gadermann, A.M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for likert rating scales. *Journal of Modern Applied Statistical Methods, 6*(1), 21-29.