


Creating Overall Measurements of Unidimensional Constructs for Comparative Research: A Methodological Study

Anastasia Charalampi

Department of Social Policy
School of Political Sciences
Panteion University of Social and Political Sciences
136 Syggrou Avenue, Athens 176 71, Greece
E-mail: acharalampi@panteion.gr
 <https://orcid.org/0000-0002-1524-3318>

Eva Tsouparopoulou

Department of Social Policy
School of Political Sciences
Panteion University of Social and Political Sciences
136 Syggrou Avenue, Athens 176 71, Greece
E-mail: etsouparopoulou@panteion.gr
 <https://orcid.org/0000-0002-0755-2508>

Joanna Tsiganou

The National Centre for Social Research
9 Kratinou & Athinas street, Athens 105 52, Greece
E-mail: jtsiganou@ekke.gr
 <https://orcid.org/0000-0002-4474-8606>

Catherine Michalopoulou

Department of Social Policy
School of Political Sciences
Panteion University of Social and Political Sciences
136 Syggrou Avenue, Athens 176 71, Greece
E-mail: kmichalop@gmail.com (kmichal@panteion.gr)
 <https://orcid.org/0000-0002-3075-7034>

Abstract: Standardization of measurement is a prerequisite for cross-national and/or overtime comparative analyses. However, there are instances in the literature where the validation of constructs resulted in producing scales or subscales defined differently from the proposed theoretical structure and across countries. In this paper, we propose an empirical methodology that provides standardized overall measurements of unidimensional constructs to be used in cross-national and overtime comparative research. Initially, the inclusion of items for further analyses is investigated at country level and overtime. The common items are to define the overall measurements and their structure is validated. Based on the Confirmatory factor analyses results, their psychometric properties are assessed. To demonstrate the implementation of the suggested methodology and facilitate practical applications, we use the human values measurements included in the European Social Survey questionnaire for Southern Europe, 2002-2018. Moreover, in order to show how these measurements may be used in further analyses, their association to subjective life satisfaction, happiness and general health are also presented.

Keywords: comparative research, confirmatory factor analysis, validity, reliability, European Social Survey

1. INTRODUCTION

Scales that have been developed as a unidimensional or multidimensional measurement of the underlying construct are extensively used in social sciences, educational, medical and health sample survey research. Scaling theory requires investigating the scales' structure and assessing their psychometric properties before their application (Michalopoulou 2017). As mentioned in our previous work (Charalampi 2018; Charalampi et al. 2019, 2020; Michalopoulou 2017), this involves splitting randomly a sample of adequate size into two halves and first performing exploratory factor analysis (EFA) on one half-sample in order to assess the construct validity of the scale. Then the structure suggested by EFA is validated by carrying out confirmatory factor analysis (CFA) on the second half-sample. Based on the full sample, the psychometric properties of the resulting scales or subscales are assessed. In the context of cross-national and/or overtime comparative research, validation of measurement should be determined first at the country level. However, there are instances in the literature where the validation of constructs resulted in providing scales or subscales defined differently from the proposed theoretical structure and across countries (see e.g. Charalampi 2018; Charalampi et al. 2020; Ertanir et al. 2021). This poses a serious methodological problem as standardization of measurement is a prerequisite for carrying out cross-national and/or overtime comparisons (Carey 2000; Kish 1994).

In this paper, we propose an empirical methodology that renders standardized overall measurements of unidimensional constructs to be used in cross-national and overtime comparative research. Initially, the inclusion of items for further

analyses is investigated at country level and overtime. The most common items are to define the overall measurements and subsequently their structure is validated. Based on the CFA results, their psychometric properties are assessed. To demonstrate the implementation of this methodology and facilitate practical applications, we use the human values measurements included in the European Social Survey (ESS) supplementary questionnaire for Southern Europe and all rounds of the ESS (2002–2018). These measurements serve as a good example because they may be defined by theory as unidimensional and are provided both at country level and overtime. Moreover, a cross-national and overtime mean scores comparison of the overall measurements of human values with satisfaction with life, happiness and general health is illustrated.

At the beginning it is worth presenting briefly the ESS human values measurements. Initially, a 57-item questionnaire was developed (Schwartz 2012; Lilleoja and Saris 2014, 2015) based on Schwartz's (1992) theory of basic human values. An alternative to this instrument was the 40-item portrait value questionnaire (PVQ), which included short verbal portraits of 40 different people. The ESS human values measurement was derived from that earlier 40-item PVQ, but because of space limitations, the number was reduced to 21 items (PVQ-21) (Davidov et al. 2008). The theory includes the following ten basic values depicted on a circular structure (Figure 1): power, achievement, hedonism, stimulation, self-direction, universalism, benevolence, tradition, conformity and security.

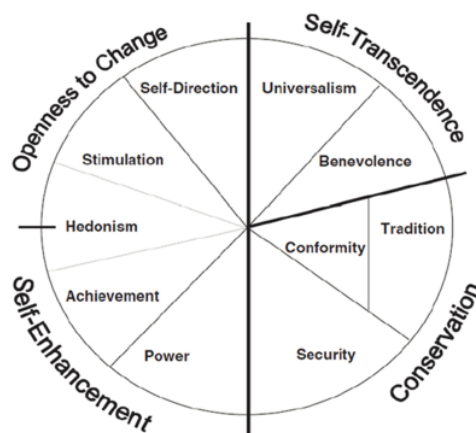


Figure 1: Structural relations among the ten values and the two dimensions. Reproduced from “Bringing values back in: the adequacy of the European Social Survey to measure values in 20 countries,” by E. Davidov, P. Schmidt and S. H. Schwartz, 2008, *Public Opinion Quarterly*, 72(3): 425. Copyright 2008 by Oxford Journals.

security (Sagin and Schwartz 2000; Davidov et al. 2008; Schwartz 2012). The circular structure also summarizes the oppositions between competing values: the self-enhancement (power and achievement) versus the self-transcendence (universalism and benevolence) dimension and the openness to change (self-direction and stimulation) versus the conservation (security, conformity and traditional) dimension (Schwartz 2012:8); hedonism shares elements of both openness to change and self-enhancement dimensions (Davidov et al. 2008; European Social Survey n.d.; Cieciuch and Davidov 2012).

Davidov (2008) and Davidov et al. (2008), investigating their cross-national comparability, found at country level that there were at least two pairs of values which were dependent on each other that could not be separated. That is why they decided to unify the strongly associated values in order to solve the problem of non-positive definite covariance matrices of the measurements. However, the definition of these unified values differed for each country and they were difficult to interpret. The current paper aims at a cross-national and overtime comparison of Schwartz's human values measurements. A common structure of each "higher order" value – openness to change, self-transcendence, conservation and self-enhancement – is created as an overall unidimensional measurement and its psychometric properties are investigated for Greece, Italy, Portugal and Spain. To decide on the inclusion of items in further analyses, item analysis was carried out using IBM SPSS Statistics Version 20. Then the resulting overall measurements were validated by performing CFA on the full samples using IBM SPSS Amos Version 21. Based on the CFA results, their psychometric properties were assessed. In order to show how these measurements may be used in further analyses, their association to subjective life satisfaction, happiness and general health are also presented.

2. THE PROPOSED METHODOLOGY: THE CASE OF UNIDIMENSIONAL CONSTRUCTS

2.1 Preliminaries

Preliminary decisions, tests and procedures are required to determine the level of measurement of scale items and the adequacy of the sample size for performing factor analyses, to investigate items' distributional properties, inspect scale items for common method variance (CMV) and decide how missing values, outliers and unengaged responses are to be dealt with.

Scale items' level of measurement. As mentioned in our previous work (Charalampi 2018; Charalampi et al. 2019), in carrying out any statistical analysis, whether it is univariate, bivariate or multivariate, the first and more important

consideration is to ascertain the level of measurement of the constructs' defining items in order to decide on the appropriateness of the methods to be used. In this respect, we are considering the case of items defined as pseudo-interval, i.e. having at least five response categories (Bartholomew et al. 2008).

Sample size adequacy. A sample size of 300 cases or more is considered adequate for performing factor analysis (Tabachnick and Fidell 2007).

Item analysis. Item analysis is performed to investigate the items' distributional properties (testing for normality) and decide on the appropriate estimation method to be used in factor analyses.

Inspecting scale items for CMV. Harman's single-factor test is performed to detect CMV. "This method loads all items from each of the constructs into an exploratory factor analysis to see whether one single factor does emerge or whether one general factor does account for a majority of the covariance between the measures; if not, the claim is that CMV is not a pervasive issue" (Chang et al. 2010:180).

Missing values. Missing data analysis is performed as complete data sets are required for SPSS Amos following the procedure described by Michalopoulou (2017).

Data screening for outliers and unengaged responses. Data screening for unengaged responses is performed and cases are to be eliminated if they exhibit low standard deviation (< 0.5), i.e. no variance in the responses (unengaged responses). Data screening for outliers is based on background variables e.g., gender (dichotomy), age (ratio) and level of educational attainment (pseudo-interval). Cases are eliminated if they are shown in the boxplots as outliers (Brown 2015; Tabachnick and Fidell 2007; Thompson 2005).

2.2 Definition of the overall measurements: Item selection

Since items are considered as pseudo-interval, the corrected item-total correlations are computed for each country and year. Internal construct validity is assessed using the criterion of corrected item-total correlations > 0.30 (Nunnally and Bernstein 1994) as the cut-off for adequate correlation (Chin et al. 2015) to decide which items to include in the analysis (Clark and Watson 1995). The most common items satisfying this criterion cross-nationally and overtime are to define the overall measurements of unidimensional constructs bearing in mind that they should be comprised by at least four items (see CFA step 1). Theoretical and empirical considerations decide on the labelling of these overall measurements.

2.3 Validation of the overall measurements' structure

Since the case of unidimensional constructs is under consideration, the structure of the overall measurements is validated by performing CFA on the full sample for each country and year.

In performing CFA on the full samples for each country and year, the following sequence of decisions is required:

1. Structure: Each overall measurement is to be considered as defining a unidimensional structure bearing in mind that “factors that are represented by two or three indicators may be underdetermined [...] and highly unstable across replications” (Brown 2015:21).
2. Model estimation: As the case of pseudo-interval items is under consideration, CFA is usually performed using the covariance matrix of association coefficients and maximum likelihood for estimation, depending on the items' distributional properties.
3. Model evaluation: Model fit is to be considered as adequate if $\chi^2/df < 3$, Standardized root-mean-square residual (SRMR) < 0.05 , Comparative fit index (CFI) and Tucker-Lewis index (TLI) values are greater than or close to 0.95 and the Root-mean-square error approximation (RMSEA) ≤ 0.06 with the 90% Confidence interval (CI) upper limit ≤ 0.06 (Bollen 1989; Brown 2015; Hu and Bentler 1999; Schmitt 2011). Model fit is considered acceptable if $\chi^2/df < 3$, SRMR < 0.08 , CFI and TLI values are > 0.90 and RMSEA < 0.08 with the 90% CI upper limit < 0.08 (Hu and Bentler 1999; Marsh et al. 2004).
4. Model misspecification searches: Searches for modification indices and further specifications are performed bearing in mind that they should not be used to overfit the model (Brown 2015). Where necessary, correlations between error variances are introduced with the required caution (Brown 2015).

2.4 Overall measurements' construction and psychometric properties assessment

The overall measurements are constructed by averaging their defining items and descriptive statistics are computed for each country and year.

Reliability assessment. Both Cronbach's alpha and Composite reliability coefficients (Raykov 2007) of the overall measurements are to be estimated. An overall measurement is considered reliable if the Composite reliability coefficient is above or around 0.70, i.e. using the same Nunnally and Bernstein (1994) criterion as for Cronbach's alpha coefficients.

Convergent validity. Based on the CFA results, the average variance extracted (AVE) is computed for each overall measurement. Convergent validity is considered

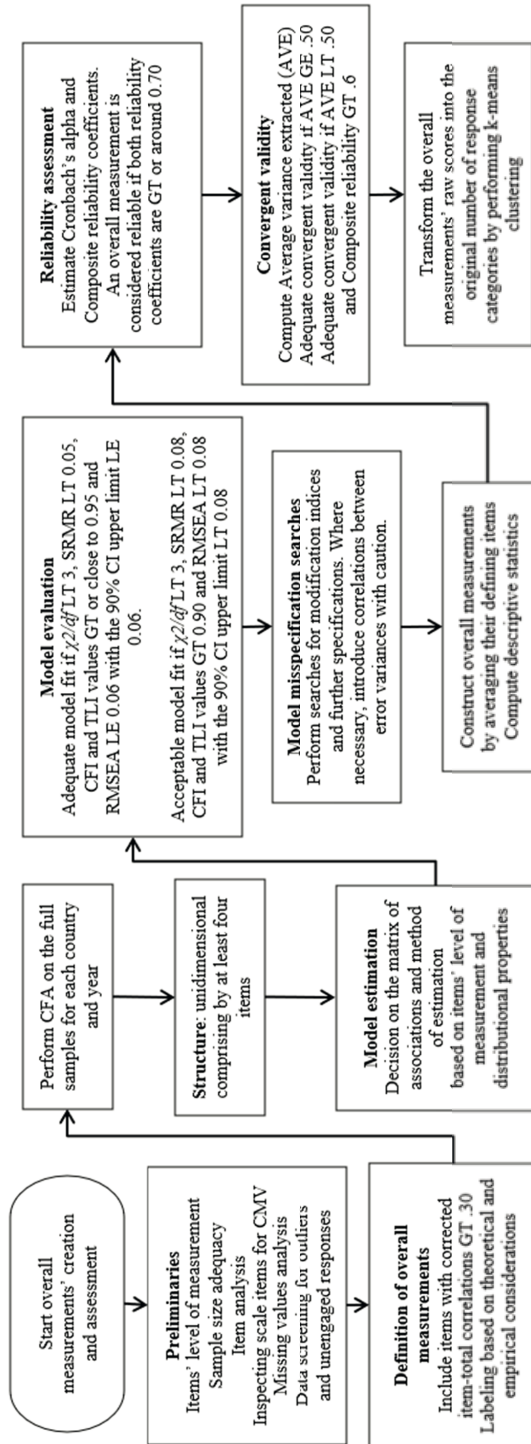


Figure 2: Process flowchart of creating and assessing the psychometric properties of overall measurements of unidimensional constructs

adequate if the AVE is above or around 0.50 (Fornell and Larcker 1981). However, if AVE is less than 0.5, but composite reliability is higher than 0.6, the convergent validity of the overall measurement is still adequate (Fornell and Larcker 1981).

Transforming the overall measurements' raw scores. In order to obtain a more meaningful representation of the respondents' scores and for better interpretation, k-means clustering is performed to transform the overall measurements' raw scores into the original number of response categories following the procedure presented by Michalopoulou and Symeonaki (2017). Cross-tabulations between the overall measurements' raw scores and the resulting clusters are performed so as to indicate the cut-off points for recoding their scores.

Illustration of the overall measurements' use in analyses. In order to illustrate how these overall measurements may be used in comparative analyses, their association to relevant scales from the literature is investigated cross-nationally and overtime. Only results statistically significant at $p < 0.05$ are to be presented.

In Figure 2, the methodological process for creating and validating overall measurements of unidimensional constructs is presented.

3. DEMONSTRATION OF THE METHODOLOGY

3.1 Participants

The analysis was based on the European Social Survey Round 1 to Round 9 Data (2002-2018) for the following countries: Greece, Italy, Portugal and Spain. Since not all four countries of our interest participated to all ESS rounds, it was decided to include in the analysis the rounds where at least three countries had participated. Therefore, the datasets of Rounds 3 (2006) and 7 (2014) were not included because only Portugal and Spain had participated. Although, Italy had participated in Round 1, the questions of human values were only asked to the male respondents. Therefore, Italy's Round 1 dataset was excluded from the analysis. It should be noted that all countries participated only in Round 2 (2004).

The ESS is centrally designed to implement all the strict methodological prerequisites for cross-national and overtime comparability (Carey 2000; Kish 1994) by applying probability sampling, minimum effective achieved sample sizes in all participating countries, a maximum target non-response rate of 30% (The ESS Sampling Expert Panel 2016), a common questionnaire and appropriate translation strategies (Harkness et al. 2010). Applying the same definition to all rounds, the ESS defines the survey population as all individuals aged 15+ residing within private households in each country, regardless of their nationality, citizenship or language. In this instance, as the age of majority is 18+ in all four participating countries, we considered that it provided a more meaningful criterion for our analyses.

The demographic and social characteristics of the participants aged 18+ are presented in Table A1 of the Appendix. As shown, in all samples there were more women than men, the mean age was over 45.8 years, at least 38% were in paid work, more than 66.2% had completed secondary education or lower and more than 50.4% of the participants were married.

3.2 Measures

The ESS human values measurement. The PVQ-21 questionnaire is worded according to the respondent's gender and is administered as a self-completion questionnaire after the end of the interview. Each item represents one of the ten values, constitutes a verbal portrait of 21 different people and describes person's goals, aspirations or wishes that show implicitly the importance of a value (Davidov et al. 2008). Each value is represented by two items, apart from universalism which is expressed by three items.

There are six possible response categories which are defined as follows: 1 (very much like me), 2 (like me), 3 (somewhat like me), 4 (a little like me), 5 (not like me) and 6 (not like me at all). Therefore, the items' level of measurement was considered as pseudo-interval. The total score for each respondent is calculated by averaging his or her responses on the items defining each value (Schwartz n.d.; Davidov et al. 2008).

Table 1 presents the ESS PVQ-21 questionnaire according to Schwartz's circular structure of the ten basic values (Davidov et al. 2008). As mentioned above, four "higher order" values encompass the ten basic values. However, the value of hedonism shares elements of two dimensions. In order to resolve this issue, reliability analysis (Cronbach's alpha coefficient) results indicated that HE21 and HE10 should be included in the openness to change and self-enhancement dimensions, respectively.

Other measures. For the purposes of our analyses, the subjective satisfaction with life, general happiness and general health scales were also used. The satisfaction with life question in the ESS questionnaire is worded as follows: "All things considered, how satisfied are you with your life as a whole nowadays? on a scale 0-10 where 0 means extremely dissatisfied and 10 means extremely satisfied". The question on happiness is worded as follows: "Taking all things together, how happy would you say you are? on a scale 0-10 where 0 means extremely unhappy and 10 means extremely happy". Finally, the general health question is worded as "How is your health in general?" taking values: 1 = very good, 2 = good, 3 = fair, 4 = bad or 5 = very bad". These values were reversed before the analysis so as to achieve correspondence between scores to the other two scales. As all three scales have at least five response categories, their level of measurement was considered as pseudo-interval.

Table 1: The European Social Survey (ESS) measurement of Schwartz's human values

Dimensions/Items	ESS supplementary questionnaire		Aligned scale	Item label
	2002, 2008	2004, 2010, 2016		
Openness to Change				
Self-Direction (SD)				
Thinking up new ideas and being creative is important to him. He likes to do things in his own original way.	Section G - A	Section H - A	1-6	SD1
It is important to him to make his own decisions about what he does. He likes to be free and not depend on others.	Section G - K	Section H - K	1-6	SD11
Stimulation (ST)				
He likes surprises and is always looking for new things to do. He thinks it is important to do lots of different things in life.	Section G - F	Section H - F	1-6	ST6
He looks for adventure and likes to take risks. He wants to have an exciting life.	Section G - O	Section H - O	1-6	ST15
Hedonism (HE)				
He seeks every chance he can to have fun. It is important to him to do things that give him pleasure.	Section G - U	Section H - U	1-6	HE21
Self-Transcendence				
Universalism (UN)				
He thinks it is important that every person in the world should be treated equally. He believes everyone should have equal opportunities in life.	Section G - C	Section H - C	1-6	UN3
It is important to him to listen to people who are different from him. Even when he disagrees with them, he still wants to understand them.	Section G - H	Section H - H	1-6	UN8
He strongly believes that people should care for nature. Looking after the environment is important to him.	Section G - S	Section H - S	1-6	UN19
Benevolence (BE)				
It's very important to him to help the people around him. He wants to care for their well-being.	Section G - L	Section H - L	1-6	BE12
It is important to him to be loyal to his friends. He wants to devote himself to people close to him.	Section G - R	Section H - R	1-6	BE18

Table 1 (continued)

Dimensions/Items	ESS supplementary questionnaire		Item label
	2002, 2008	2004, 2010, 2016, 2018	
Conservation			
Tradition (TR)			
It is important to him to be humble and modest. He tries not to draw attention to himself.	Section G - I	Section H - I	1-6 TR9
Tradition is important to him. He tries to follow the customs handed down by his religion or his family.	Section G - T	Section H - T	1-6 TR20
Conformity (CO)			
He believes that people should do what they're told. He thinks people should follow rules at all time, even when no-one is watching.	Section G - G	Section H - G	1-6 CO7
It is important to him always to behave properly. He wants to avoid doing anything people would say is wrong.	Section G - P	Section H - P	1-6 CO16
Security (SEC)			
It is important to him to live in secure surroundings. He avoids anything that might endanger his safety.	Section G - E	Section H - E	1-6 SEC5
It is important to him that the government ensures his safety against all threats. He wants the state to be			
strong so it can defend its citizens.	Section G - N	Section H - N	1-6 SEC14
Self-Enhancement			
Power (PO)			
It is important to him to be rich. He wants to have a lot of money and expensive things.	Section G - B	Section H - B	1-6 PO2
It is important to him to get respect from others. He wants people to do what he says.	Section G - Q	Section H - Q	1-6 PO17
Achievement (AC)			
It's important to him to show his abilities. He wants people to admire what he does.	Section G - D	Section H - D	1-6 AC4
Being very successful is important to him. He hopes people will recognize his achievements.	Section G - M	Section H - M	1-6 AC13
Hedonism (HE)			
Having a good time is important to him. He likes to "spoil" himself.	Section G - J	Section H - J	1-6 HE10

The wording of the ESS human values questions is as follows: "Here we briefly describe some people. Please read each description and tick the box on each line that shows how much each person is or is not like you". 1 = very much like me, 2 = like me, 3 = somewhat like me, 4 = a little like me, 5 = not like me, 6 = not like me at all. HE or SHE. The value of hedonism shares elements of both openness to change and self-enhancement. Cronbach's alpha coefficient provided better results by including HE21 and HE10 in the openness to change and self-enhancement dimensions, respectively.

3.3 Preliminaries

Preliminary decisions, tests and procedures were carried out pertaining to the level of measurement of scale items and the adequacy of the sample size for performing factor analyses, the investigation of items' distributional properties, inspecting scale items for CMV and deciding how missing values, outliers and unengaged responses are to be dealt with.

Scale items' level of measurement. As mentioned above, all the human values scale items were considered as pseudo-interval.

Sample size adequacy. The sample sizes ranged from 935 (Italy, 2012) to 2,685 (Italy, 2018) and therefore were considered large enough (> 300) for carrying out factor analyses separately in each country and year.

Item analysis. In all cases, item analyses revealed that non-normality was not severe (skewness >2 ; kurtosis >7) for any item (West et al. 1995).

Inspecting scale items for CMV. Harman's single-factor test was performed to detect CMV using Principal axis factoring as the factor extraction method (Table A2 of the Appendix). As the percent of the total variance explained ranged from 18.989 (Spain, 2016) to 33.816 (Portugal, 2010), we concluded that CMV was not "a pervasive issue" (Chang et al. 2010:180).

Missing values. Missing data ranged from 0.1 percent for the items BE12, BE18, TR20, SEC14, PO2 and PO17 (Greece, 2004) to 4.3 percent for HE10 (Italy, 2004) and they were dealt with regression imputation given the limitations of the software used (Michalopoulou 2017).

Data screening for outliers and unengaged responses. Screening the datasets identified only a few unengaged responses (standard deviation = .000) and outlying cases with higher education degree in each dataset and it was decided not to reject them from the analysis.

3.4 Definition of the overall measurements: Item selection

In Table 2, the corrected item-total correlations are presented for each country and year. As shown, one item of openness to change (ST15), self-transcendence (UN3) and self-enhancement (PO2) were below the criterion of corrected item-total correlations $> .30$ in many cases. Accordingly, up to five out of the six items of conservation were below this limit in certain cases and this "higher order" value was excluded from further analysis. Therefore, three overall measurements with comparable structures cross-nationally and overtime consisting of four items each remained in the analysis. Since only one item was excluded from their respective "higher order" value, it was decided to retain the corresponding dimensional label: openness to change, self-transcendence and self-enhancement.

Table 2: Corrected item-total correlations of Schwartz's scale values for Southern Europe: European Social Survey, 2002–2018

Dimension/ Item	Country																												
	Greece						Italy						Portugal						Spain										
	2002	2004	2008	2010	2004	2012	2016	2018	2002	2004	2008	2010	2012	2016	2018	2002	2004	2008	2010	2012	2016	2018	2002	2004	2008	2010	2012	2016	2018
N	2,511	2,363	2,019	2,649	1,497	935	2,524	2,685	1,444	2,008	2,296	2,104	2,117	1,249	1,034	1,670	1,608	2,486	1,840	1,841	1,918	1,613							
Openness to change																													
SD1	0.481	0.545	0.520	0.518	0.446	0.402	0.495	0.486	0.436	0.640	0.486	0.527	0.347	0.343	0.379	0.482	0.336	0.366	0.331	0.527	0.317	0.304							
SD11	0.540	0.568	0.498	0.431	0.430	0.401	0.519	0.544	0.560	0.643	0.524	0.563	0.402	0.434	0.434	0.530	0.493	0.412	0.311	0.590	0.327	0.372							
ST6	0.529	0.526	0.502	0.468	0.497	0.485	0.540	0.518	0.513	0.508	0.385	0.545	0.412	0.451	0.415	0.467	0.413	0.471	0.437	0.520	0.426	0.455							
ST15	0.349	0.357	0.353	0.329	0.278	0.231	0.218	0.268	0.405	0.296	0.231	0.279	0.288	0.337	0.277	0.272	0.257	0.377	0.322	0.293	0.327	0.345							
HE21	0.510	0.537	0.479	0.445	0.475	0.408	0.421	0.425	0.399	0.393	0.385	0.430	0.396	0.408	0.462	0.486	0.404	0.494	0.412	0.411	0.454	0.457							
Self-transcendence																													
UN3	0.396	0.404	0.407	0.318	0.236	0.321	0.423	0.426	0.437	0.520	0.501	0.495	0.278	0.398	0.451	0.412	0.432	0.276	0.213	0.442	0.224	0.273							
UN8	0.524	0.538	0.520	0.441	0.367	0.329	0.514	0.480	0.539	0.616	0.568	0.628	0.356	0.417	0.443	0.484	0.461	0.424	0.311	0.546	0.303	0.326							
UN19	0.490	0.489	0.501	0.439	0.347	0.343	0.507	0.563	0.526	0.591	0.576	0.607	0.321	0.485	0.513	0.492	0.455	0.327	0.317	0.544	0.334	0.315							
BE12	0.423	0.418	0.487	0.393	0.399	0.473	0.536	0.559	0.579	0.574	0.571	0.609	0.332	0.486	0.533	0.531	0.499	0.345	0.294	0.518	0.342	0.350							
BE18	0.435	0.534	0.507	0.378	0.414	0.476	0.534	0.600	0.581	0.606	0.566	0.569	0.413	0.511	0.532	0.564	0.476	0.382	0.357	0.531	0.376	0.377							
Conservation																													
TR9	0.147	0.194	0.270	0.204	0.164	0.216	0.388	0.363	0.324	0.386	0.423	0.493	0.227	0.346	0.382	0.329	0.318	0.107	0.059	0.363	0.229	0.225							
TR20	0.233	0.265	0.329	0.258	0.197	0.262	0.335	0.348	0.250	0.396	0.429	0.433	0.208	0.285	0.283	0.221	0.243	0.067	0.133	0.337	0.228	0.256							
CO7	0.240	0.286	0.347	0.185	0.221	0.233	0.307	0.273	0.253	0.344	0.388	0.403	0.283	0.232	0.198	0.210	0.177	0.197	0.202	0.348	0.178	0.193							
CO16	0.393	0.482	0.475	0.317	0.315	0.334	0.454	0.493	0.328	0.356	0.375	0.479	0.328	0.306	0.284	0.431	0.408	0.228	0.265	0.461	0.268	0.314							
SEC5	0.388	0.383	0.416	0.348	0.344	0.394	0.450	0.504	0.435	0.544	0.509	0.561	0.374	0.412	0.458	0.425	0.437	0.292	0.286	0.568	0.318	0.378							
SEC14	0.440	0.446	0.455	0.345	0.332	0.448	0.461	0.509	0.529	0.577	0.456	0.499	0.362	0.449	0.451	0.425	0.397	0.341	0.309	0.519	0.314	0.370							
Self-enhancement																													
PO2	0.317	0.391	0.437	0.360	0.277	0.291	0.224	0.238	0.291	0.386	0.220	0.273	0.291	0.106	0.106	0.254	0.152	0.310	0.208	0.277	0.230	0.272							
PO17	0.392	0.442	0.494	0.430	0.397	0.370	0.444	0.449	0.388	0.459	0.452	0.534	0.341	0.342	0.322	0.396	0.356	0.333	0.336	0.523	0.318	0.316							
AC4	0.543	0.520	0.537	0.504	0.441	0.484	0.481	0.506	0.479	0.586	0.591	0.623	0.445	0.464	0.401	0.474	0.449	0.417	0.422	0.601	0.399	0.426							
AC13	0.550	0.567	0.553	0.527	0.518	0.591	0.643	0.652	0.586	0.618	0.587	0.610	0.485	0.473	0.505	0.530	0.468	0.467	0.459	0.550	0.488	0.495							
HE10	0.469	0.540	0.532	0.413	0.425	0.355	0.363	0.428	0.588	0.624	0.535	0.551	0.404	0.462	0.524	0.526	0.449	0.504	0.425	0.539	0.448	0.475							

Corrected item-total correlations < 0.30 are in boldface.

3.5 Validation of the overall measurements of human values structure

The common structures of the three overall measurements (openness to change, self-transcendence and self-enhancement) were investigated by performing CFA on the full samples for each country and year (Table 3). As shown, these analyses produced models with adequate model fit for 33 cases and acceptable model fit for 23 cases. In the following 10 cases model fit was inadequate: Italy (self-transcendence, 2016, 2018), Portugal (openness to change, 2002, 2018; self-transcendence, 2018) and Spain (openness to change, 2002, 2016; self-transcendence, 2002; self-enhancement; 2002, 2004).

The standardized solutions for the first-order factors of the three overall measurements (openness to change, self-transcendence and self-enhancement) based on CFA analysis performed on the full samples of the four Southern European countries are presented cross-nationally in Figure 3. Round 2 (2004) was selected as an example to be included in the main text as this is the only round with data available from all four countries.

Table 3: Confirmatory factor analysis with maximum likelihood of the covariance matrix performed on the full samples of the Southern European countries, goodness-of-fit indices: European Social Survey, 2002–2018

	χ^2/df	SRMR	CFI	TLI	RMSEA(90% CI)
Greece					
2002 (<i>N</i> = 2,511)					
Openness to change	0.72	0.003	1.000	1.001	0.001 (0.001-0.050)
Self-transcendence	5.49	0.009	0.998	0.994	0.026 (0.001-0.054)
Self-enhancement	2.33	0.006	0.999	0.996	0.023 (0.001-0.063)
2004 (<i>N</i> = 2,363)					
Openness to change	0.78	0.003	1.000	1.001	0.001 (0.001-0.052)
Self-transcendence	0.23	0.002	1.000	1.002	0.001 (0.001-0.019)
Self-enhancement	0.26	0.002	1.000	1.002	0.001 (0.001-0.044)
2008 (<i>N</i> = 2,019)					
Openness to change	2.96	0.008	0.999	0.993	0.031 (0.001-0.075)
Self-transcendence	0.23	0.006	1.000	1.004	0.001 (0.001-0.030)
Self-enhancement	0.54	0.003	1.000	1.002	0.001 (0.001-0.053)
2010 (<i>N</i> = 2,649)					
Openness to change	0.50	0.003	1.001	1.000	0.001 (0.001-0.045)
Self-transcendence	1.68	0.005	1.000	0.997	0.016 (0.001-0.057)
Self-enhancement	4.54	0.010	0.998	0.995	0.022 (0.001-0.049)

	χ^2/df	SRMR	CFI	TLI	RMSEA (90% CI)
Italy					
2004 (N = 1,497)					
Openness to change	6.57	0.014	0.995	0.984	0.039 (0.008-0.074)
Self-transcendence	4.78	0.011	0.997	0.992	0.030 (0.001-0.067)
Self-enhancement	1.94	0.008	1.000	1.000	0.001 (0.001-0.051)
2012 (N = 935)					
Openness to change	2.14	0.011	1.000	0.999	0.009 (0.001-0.066)
Self-transcendence	0.01	0.001	1.000	1.011	0.001 (0.001-0.001)
Self-enhancement	2.84	0.011	0.999	0.996	0.021 (0.001-0.072)
2016 (N = 2,524)					
Openness to change	1.91	0.006	0.999	0.997	0.019 (0.001-0.060)
Self-transcendence	12.87	0.011	0.996	0.974	0.069 (0.039-0.104)
Self-enhancement	1.99	0.006	0.999	0.996	0.020 (0.001-0.061)
2018 (N = 2,685)					
Openness to change	1.80	0.006	1.000	0.998	0.017 (0.001-0.058)
Self-transcendence	10.95	0.010	0.997	0.982	0.061 (0.032-0.096)
Self-enhancement	4.53	0.009	0.999	0.996	0.022 (0.001-0.049)
Portugal					
2002 (N = 1,444)					
Openness to change	15.28	0.018	0.988	0.927	0.099 (0.060-0.146)
Self-transcendence	6.85	0.011	0.997	0.992	0.041 (0.010-0.076)
Self-enhancement	1.55	0.007	1.000	0.997	0.020 (0.001-0.076)
2004 (N = 2,008)					
Openness to change	3.66	0.007	0.999	0.993	0.036 (0.001-0.079)
Self-transcendence	2.71	0.006	0.999	0.996	0.029 (0.001-0.073)
Self-enhancement	2.89	0.008	0.999	0.994	0.031 (0.001-0.075)
2008 (N = 2,296)					
Openness to change	1.88	0.005	0.999	0.996	0.020 (0.001-0.063)
Self-transcendence	0.31	0.002	1.000	1.001	0.001 (0.001-0.045)
Self-enhancement	1.95	0.006	0.999	0.997	0.020 (0.001-0.063)
2010 (N = 2,104)					
Openness to change	0.01	0.001	1.000	1.004	0.001 (0.001-0.001)
Self-transcendence	0.21	0.002	1.000	1.005	0.001 (0.001-0.048)
Self-enhancement	0.76	0.004	1.000	1.001	0.001 (0.001-0.050)
2012 (N = 2,117)					
Openness to change	3.83	0.008	0.998	0.989	0.037 (0.001-0.078)
Self-transcendence	2.94	0.007	0.999	0.995	0.030 (0.001-0.073)
Self-enhancement	0.01	0.001	1.000	1.003	0.001 (0.001-0.024)

	χ^2/df	SRMR	CFI	TLI	RMSEA (90% CI)
2016 (N = 1,249)					
Openness to change	0.87	0.006	1.000	1.001	0.001 (0.001-0.073)
Self-transcendence	2.70	0.009	0.999	0.998	0.017 (0.001-0.061)
Self-enhancement	1.69	0.008	1.000	1.001	0.001 (0.001-0.053)
2018 (N = 1,034)					
Openness to change	2.23	0.010	0.998	0.987	0.035 (0.001-0.098)
Self-transcendence	2.31	0.009	0.999	0.992	0.036 (0.001-0.098)
Self-enhancement	0.06	0.002	1.000	1.011	0.001 (0.001-0.052)
Spain					
2002 (N = 1,670)					
Openness to change	5.89	0.010	0.996	0.977	0.054 (0.019-0.100)
Self-transcendence	3.78	0.009	0.998	0.990	0.041 (0.001-0.088)
Self-enhancement	3.22	0.010	0.998	0.989	0.036 (0.001-0.084)
2004 (N = 1,608)					
Openness to change	0.63	0.004	1.000	1.003	0.001 (0.001-0.061)
Self-transcendence	7.97	0.013	0.996	0.987	0.043 (0.015-0.076)
Self-enhancement	4.02	0.012	0.997	0.980	0.043 (0.005-0.091)
2008 (N = 2,486)					
Openness to change	1.56	0.005	1.000	0.998	0.015 (0.001-0.058)
Self-transcendence	1.52	0.053	1.000	0.998	0.014 (0.001-0.058)
Self-enhancement	1.04	0.005	1.000	1.000	0.004 (0.001-0.053)
2010 (N = 1,840)					
Openness to change	1.57	0.007	0.999	0.996	0.018 (0.001-0.068)
Self-transcendence	0.21	0.002	1.000	1.005	0.001 (0.001-0.048)
Self-enhancement	1.04	0.005	1.000	1.002	0.001 (0.001-0.038)
2012 (N = 1,841)					
Openness to change	1.15	0.006	1.000	0.999	0.009 (0.001-0.063)
Self-transcendence	2.81	0.008	0.999	0.998	0.015 (0.001-0.051)
Self-enhancement	0.35	0.003	1.000	1.004	0.001 (0.001-0.052)
2016 (N = 1,918)					
Openness to change	10.98	0.015	0.990	0.943	0.072 (0.038-0.113)
Self-transcendence	2.66	0.008	0.998	0.989	0.029 (0.001-0.075)
Self-enhancement	0.34	0.003	1.000	1.003	0.001 (0.001-0.050)
2018 (N = 1,613)					
Openness to change	0.19	0.003	1.000	1.007	0.001 (0.001-0.050)
Self-transcendence	0.05	0.001	1.000	1.005	0.001 (0.001-0.040)
Self-enhancement	0.18	0.003	1.000	1.000	0.001 (0.001-0.050)

df = degrees of freedom; SRMR = standardized root mean square residual; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root-mean-square error of approximation; CI = confidence interval. Model fit is considered adequate if SRMR < 0.08, CFI > 0.90, TLI > 0.90 and RMSEA < 0.08 with the 90% CI upper limit < 0.08.

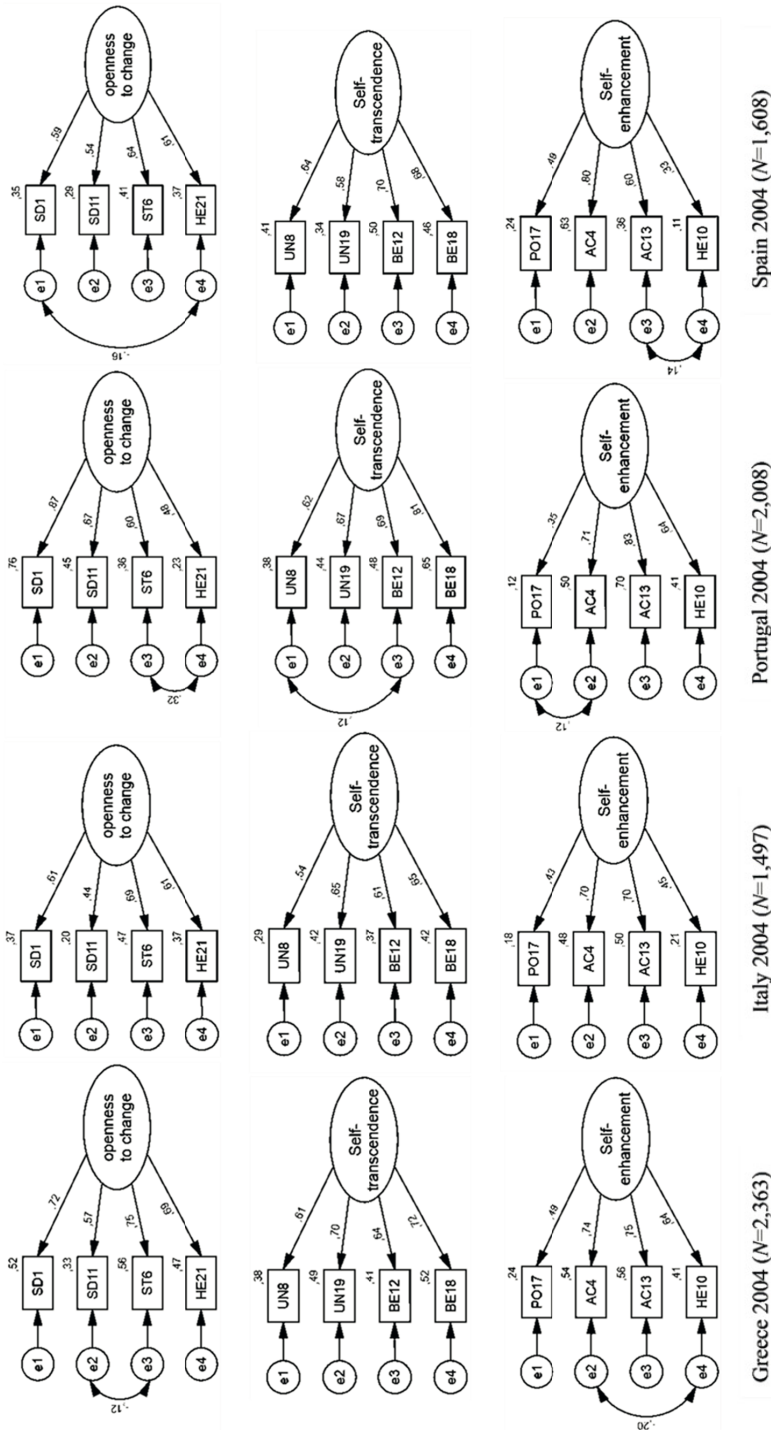


Figure 3: Standardized solutions for the first-order factors of openness to change, self-transcendence and self-enhancement based on CFA analysis performed on the full samples of the Southern European Survey, 2004. Observed and latent variables are represented by rectangles and ellipses, respectively.

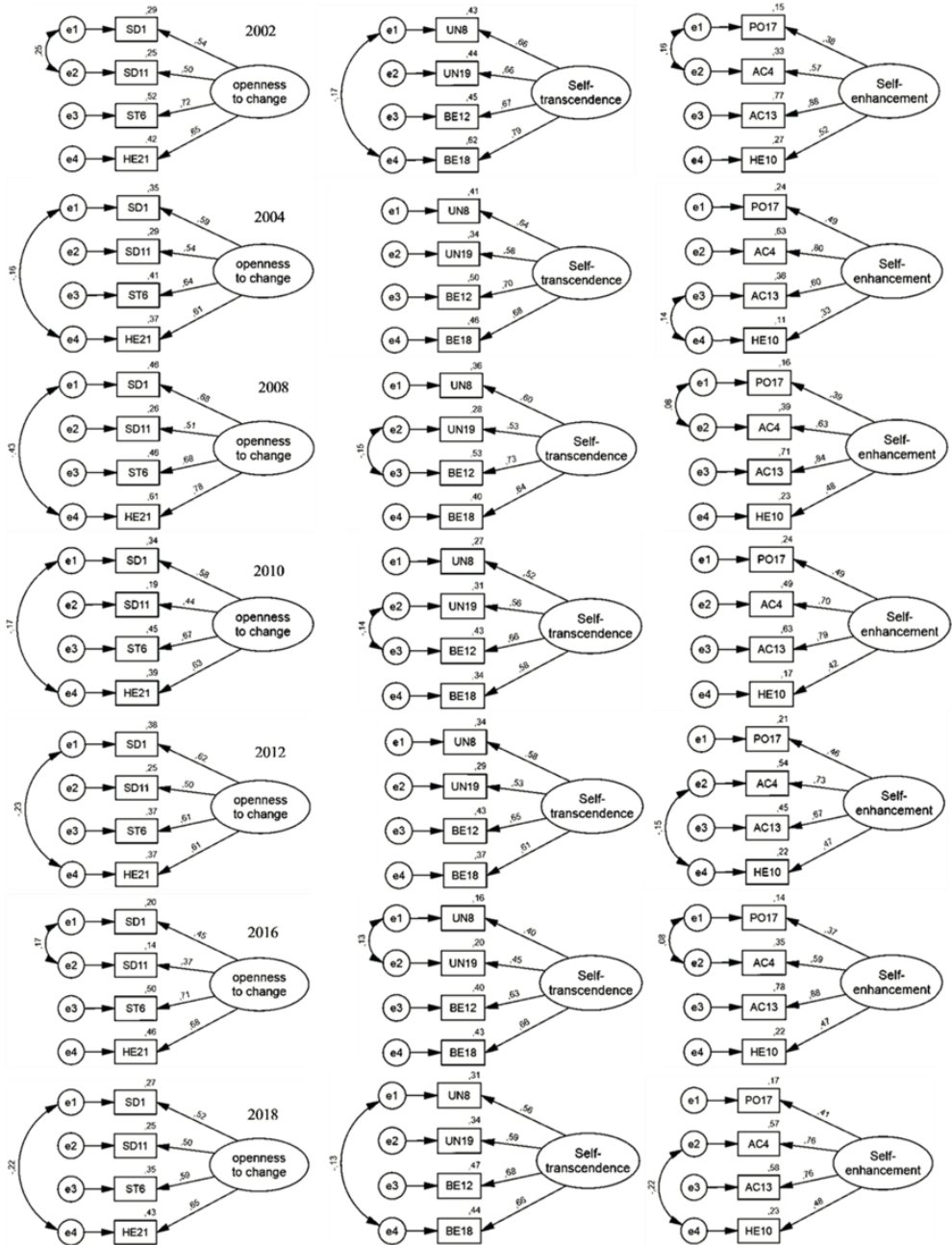


Figure 4: Standardized solutions for the first-order factors of openness to change, self-transcendence and self-enhancement based on CFA analysis performed on the full samples of Spain: European Social Survey, 2002-2018. Observed and latent variables are represented by rectangles and ellipses, respectively.

The standardized solutions for the first-order factors of the three overall measurements (openness to change, self-transcendence and self-enhancement) based on CFA analysis performed on the full samples of Spain are presented in Figure 4 for Round 1 (2002) to Round 9 (2018). Spain and Portugal have participated in all nine Rounds of the ESS and Spain was selected as an example of an overtime presentation. Because of space limitation and in order to avoid the repetition of figures by presenting the solutions both cross-nationally and overtime, it was decided to better present them only overtime in Figures A1-A3 in the Appendix for Greece, Italy and Portugal, respectively.

Table 4: Reliabilities and convergent validity of the three human values subscales for the Southern European countries: European Social Survey, 2002–2018

	Openness to change			Self-transcendence			Self-enhancement		
	CA	CR	CV	CA	CR	CV	CA	CR	CV
Greece									
2002	0.755	0.771	0.458	0.714	0.712	0.389	0.716	0.715	0.391
2004	0.768	0.779	0.469	0.760	0.763	0.448	0.730	0.754	0.439
2008	0.724	0.755	0.433	0.739	0.751	0.429	0.725	0.724	0.402
2010	0.737	0.759	0.442	0.673	0.694	0.362	0.665	0.677	0.358
Italy									
2004	0.689	0.681	0.354	0.708	0.707	0.376	0.650	0.664	0.342
2012	0.611	0.610	0.287	0.675	0.667	0.345	0.686	0.696	0.376
2016	0.706	0.728	0.406	0.777	0.768	0.460	0.660	0.692	0.372
2018	0.716	0.735	0.412	0.795	0.792	0.397	0.699	0.716	0.494
Portugal									
2002	0.716	0.710	0.383	0.790	0.798	0.501	0.713	0.722	0.418
2004	0.766	0.757	0.447	0.798	0.793	0.488	0.728	0.738	0.432
2008	0.668	0.646	0.329	0.826	0.828	0.551	0.698	0.707	0.391
2010	0.695	0.676	0.358	0.811	0.671	0.337	0.726	0.720	0.401
2012	0.712	0.701	0.379	0.755	0.752	0.443	0.715	0.718	0.403
2016	0.628	0.663	0.331	0.726	0.742	0.424	0.642	0.657	0.340
2018	0.661	0.697	0.368	0.742	0.772	0.460	0.627	0.667	0.354
Spain									
2002	0.717	0.698	0.370	0.778	0.790	0.485	0.692	0.690	0.380
2004	0.670	0.687	0.354	0.751	0.746	0.426	0.648	0.650	0.335
2008	0.724	0.761	0.444	0.694	0.721	0.393	0.671	0.685	0.371
2010	0.650	0.672	0.345	0.646	0.671	0.337	0.687	0.700	0.384
2012	0.644	0.676	0.344	0.678	0.685	0.356	0.652	0.677	0.354
2016	0.647	0.644	0.325	0.622	0.620	0.297	0.667	0.679	0.371
2018	0.626	0.653	0.322	0.700	0.718	0.387	0.675	0.704	0.392

CA = Cronbach's alpha; CR = composite reliability; CV = convergent validity. Reliability coefficients ≥ 0.70 and convergent validity ≥ 0.50 are in boldface.

3.6 Overall measurements of human values construction and assessment

The three overall measurements of human values (openness to change, self-transcendence and self-enhancement) were constructed by averaging their defining items based on the full samples of each country and year. In Table 4, Cronbach's alpha and composite reliability coefficients and convergent validity of the overall measurements are presented for each country and year. As shown, Cronbach's alpha was above or around 0.70 in 56 out of 66 overall measurements, whereas composite reliability in 62. Based on composite reliability, only four cases were below that limit: Italy (openness to change, 2012), Portugal (openness to change, 2008) and Spain (openness to change, 2016; self-transcendence, 2016). As reliability coefficients were higher than 0.6, the convergent validity of the overall measurements was considered adequate in all cases.

In order to obtain a more meaningful representation of the respondents' scores and for better interpretation, k-means clustering (k=6) was performed and the overall measurements' raw scores were transformed into the original six response categories.

3.7 Overall measurements of human values on satisfaction with life, happiness and general health: Cross-national mean scores comparison

In Figure 5, a cross-national demonstration of mean scores comparison of the three overall measurements of human values (openness to change, self-transcendence and self-enhancement) with satisfaction with life, happiness and general health is presented for the ESS Round 2 (2004).

In the case of Greece, higher levels (1) of openness to change, self-transcendence and self-enhancement were obtained for those inclined to the higher end (10) of satisfaction with life scale with mean scores 2.58, 1.64 and 2.82, respectively. Higher levels (1) of openness to change, self-transcendence and self-enhancement were obtained for those inclined to the higher end (10) of happiness scale with mean scores 2.61, 1.60 and 2.71, respectively. Higher levels (1) of openness to change, self-transcendence and self-enhancement were obtained for those inclined to the higher end (5) of general health scale with mean scores 2.61, 1.69 and 2.48, respectively.

Moreover, in the case of Italy, higher levels (1) of openness to change and self-transcendence were obtained for those inclined to the higher end (10) of satisfaction with life scale with mean scores 2.40 and 1.76, respectively. Higher levels (1) of openness to change, self-transcendence and self-enhancement were obtained for those inclined to the higher end (10) of happiness scale with mean scores 2.25, 1.73 and 3.15, respectively. Higher levels (1) of openness to change

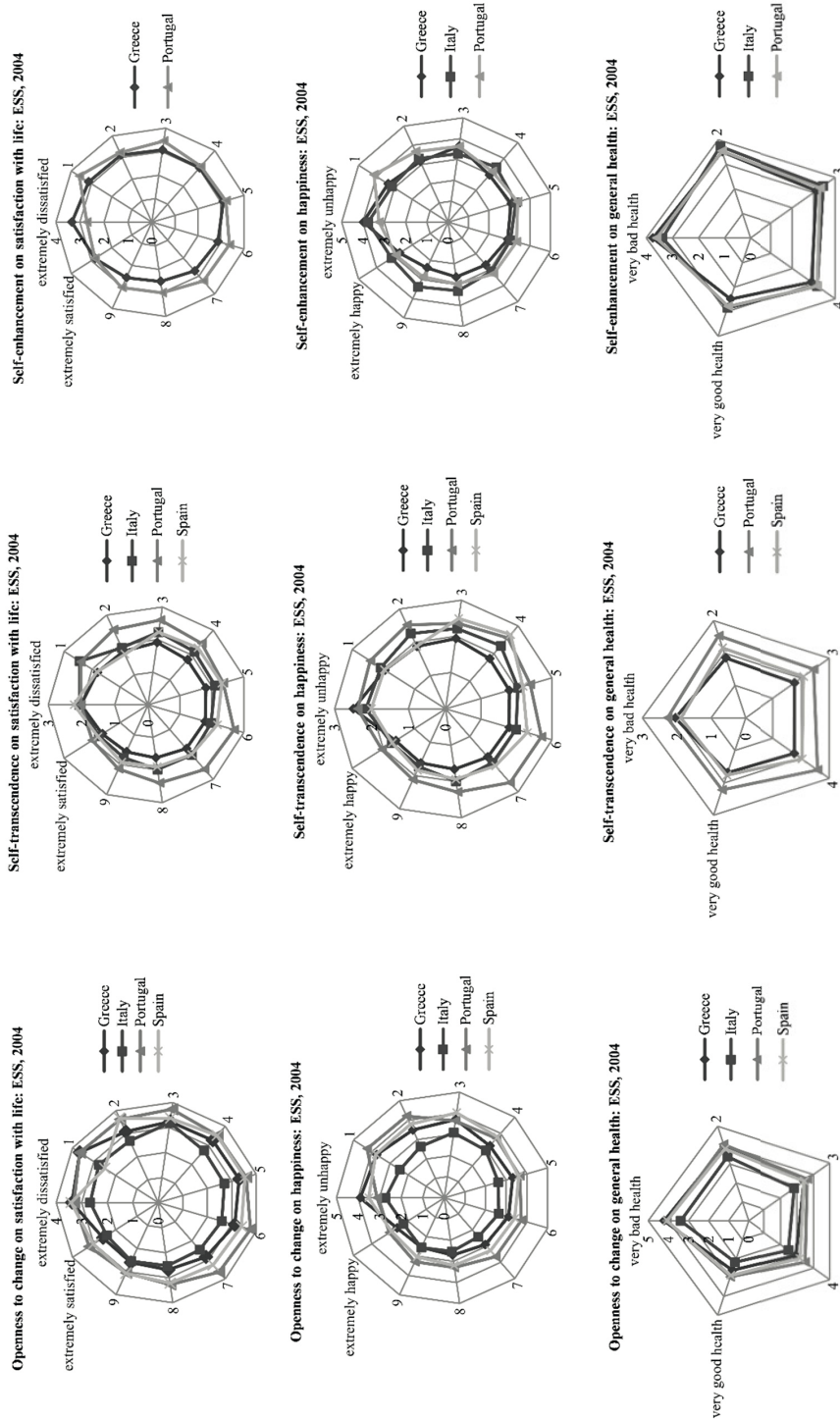


Figure 5. The mean scores of openness to change, self-transcendence and self-enhancement on satisfaction with life, happiness and general health scale for the Southern European countries: European Social Survey (ESS), 2004. Data weighted by: dweight (design weight) * pweight (population size weight).

and self-enhancement were obtained for those inclined to the higher end (5) of general health scale with mean scores 2.19 and 2.84, respectively.

In the case of Portugal, higher levels (1) of openness to change, self-transcendence and self-enhancement were obtained for those inclined to the higher end (10) of satisfaction with life scale with mean scores 3.23, 2.00 and 2.82, respectively. Higher levels (1) of openness to change, self-transcendence and self-enhancement were obtained for those inclined to the higher end (10) of happiness scale with mean scores 3.02, 2.05 and 2.91, respectively. Higher levels (1) of openness to change, self-transcendence and self-enhancement were obtained for those inclined to the higher end (5) of general health scale with mean scores 2.96, 2.22 and 2.78, respectively.

Finally in the case of Spain, higher levels (1) of openness to change and self-transcendence were obtained for those inclined to the higher end (10) of satisfaction with life scale with mean scores 3.03 and 1.79, respectively. Higher levels (1) of openness to change and self-transcendence were obtained for those inclined to the higher end (10) of happiness scale with mean scores 2.90 and 1.71, respectively. Higher levels (1) of openness to change and self-transcendence were obtained for those inclined to the higher end (5) of general health scale with mean scores 2.89 and 1.77, respectively.

Therefore, similar patterns of the overall measurements of human values and satisfaction with life, happiness and general health were observed in all countries for the ESS of 2004.

3.8 Overall measurements of human values on satisfaction with life, happiness and general health: Overtime mean scores comparison

In Figure 6, an overtime demonstration of mean scores comparison of the three overall measurements of human values (openness to change, self-transcendence and self-enhancement) to satisfaction with life, happiness and general health is presented for the ESS datasets of Spain for 2002-2018.

In the datasets of 2004, 2010 and 2012, higher levels (1) of openness to change were obtained for those inclined to the higher end (10) of satisfaction with life scale with mean scores 3.03, 2.85 and 2.54, respectively. The reverse holds true for the datasets of 2008 and 2018. For datasets of 2004, 2008, 2010, 2012 and 2018, higher levels (1) of openness to change were obtained for those inclined to the higher end (10) of happiness scale with mean scores 2.90, 2.64, 2.83, 2.43 and 2.64, respectively. For the same years, higher levels (1) of openness to change were obtained for those inclined to the higher end (5) of general health scale with mean scores 2.89, 2.59, 2.58, 2.22 and 2.51, respectively.

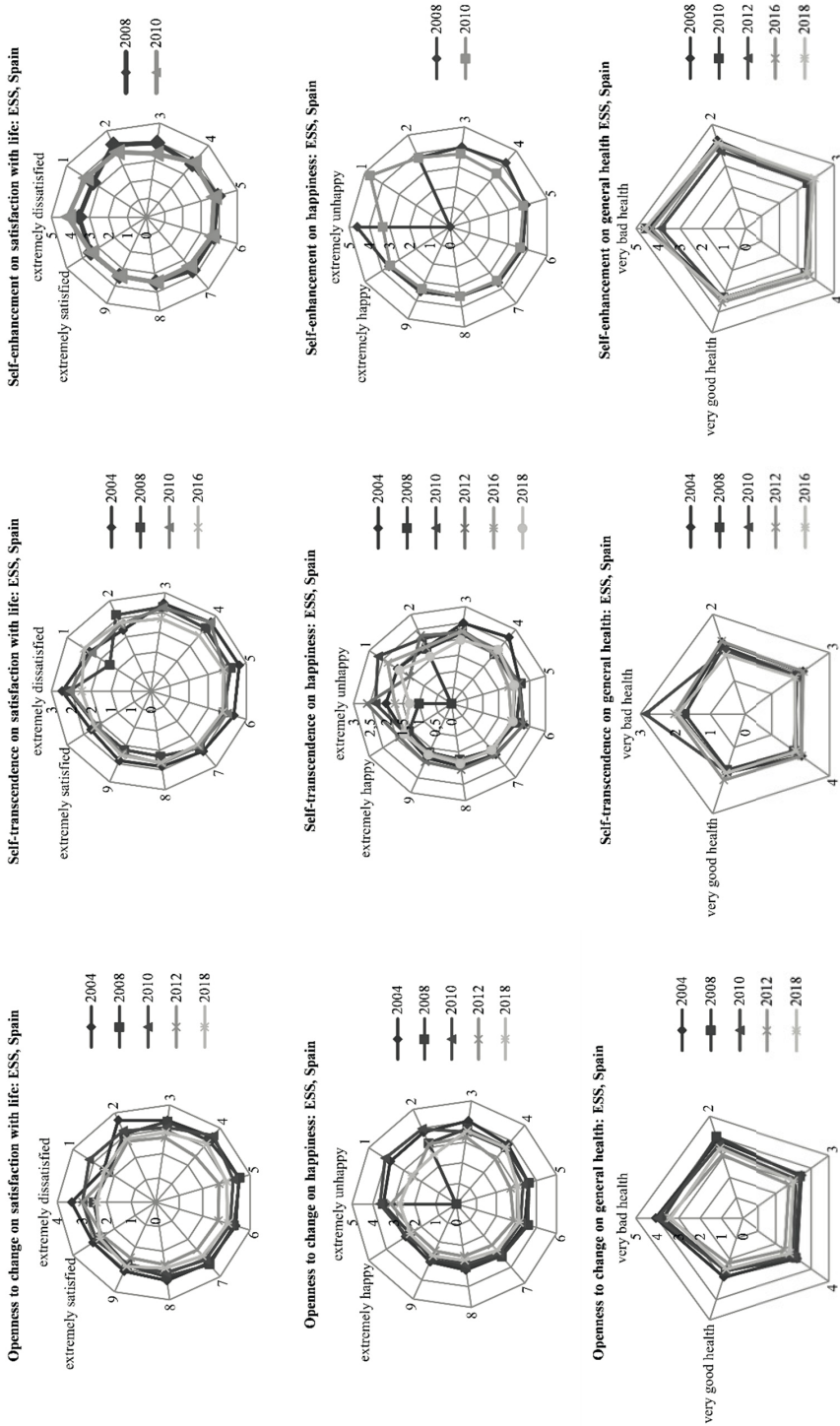


Figure 6. The mean scores of openness to change, self-transcendence and self-enhancement on satisfaction with life, happiness and general health scale for Spain: European Social Survey (ESS), 2002-2018. Data weighted by: dweight (design weight) * pweight (population size weight).

In the datasets of 2004, 2008, 2010 and 2016, higher levels (1) of self-transcendence were obtained for those inclined to the higher end (10) of satisfaction with life scale with mean scores 1.79, 1.63, 1.61 and 1.68, respectively. For datasets of 2004, 2010, 2012 and 2016, higher levels (1) of self-transcendence were obtained for those inclined to the higher end (10) of happiness scale with mean scores 1.71, 1.68, 1.90 and 1.71, respectively. The reverse holds true for the datasets of 2008 and 2018. For datasets of 2004, 2008, 2010, 2012 and 2016, higher levels (1) of self-transcendence were obtained for those inclined to the higher end (5) of general health scale with mean scores 1.77, 1.68, 1.75, 1.99 and 1.78, respectively.

In the datasets of 2008 and 2010, higher levels (1) of self-enhancement were obtained for those inclined to the higher end (10) of satisfaction with life scale with mean scores 3.48 and 3.46, respectively. For the dataset of 2008, higher levels (1) of self-enhancement were obtained for those inclined to the higher end (10) of happiness scale with mean score 3.61. The reverse holds true for the dataset of 2010. Finally, for datasets of 2008, 2010, 2012, 2016 and 2018, higher levels (1) of self-enhancement were obtained for those inclined to the higher end (5) of general health scale with mean scores 3.30, 3.30, 3.34, 3.54 and 3.33, respectively.

Therefore, a similar pattern of the overall measurements of human values and satisfaction with life, happiness and general health were observed for Spain in all Rounds (2002-2018) of the ESS but with a few exceptions.

4. DISCUSSION AND CONCLUSIONS

Cross-national and/or overtime comparative research requires the detailed investigation of all aspects of the sample survey design and its implementation as defined by Kish (1994) and determined in detail by Carey (2000). The requirement for measurement is standardization and, if this is not feasible to establish, then measurements should be harmonized by creating “a desired degree of comparability between statistics of different countries” (Ehling 2003:17) to ensure their uniformity cross-nationally and overtime (Michalopoulou 2016). However, when attitude scales are under consideration, their validation at country level has resulted in some instances at producing different structure than the theoretical one and across countries. This paper aims to present an empirical methodology that provides standardized overall measurements of unidimensional constructs to be used in cross-national and overtime comparative research. The demonstrated methodology was based on the ESS human values measurements for Greece, Italy, Portugal and Spain and all Rounds of the survey (2002–2018) since ESS is centrally designed so as to allow for cross-national and overtime comparability of measurement. Four unidimensional measurements defined as the theoretical

“higher order” values of openness to change, self-transcendence, conservation and self-enhancement were considered so as to indicate how the proposed methodology may cover also the case of multidimensional constructs. The criterion of corrected item-total correlation was employed for items’ selection and conservation was excluded from further analysis. Overall measurements comprised of four items each were constructed for the other three constructs based on the remaining items in the analysis and their validation showed that they were both reliable and valid for most of the cases and therefore may be compared cross-nationally and overtime. In order to show how these overall measurements of human values may be used in analyses, their mean scores on satisfaction with life, happiness and general health scales were investigated cross-nationally for Greece, Italy, Portugal and Spain based on the Round 2 (2004) datasets and overtime for Spain that had participated in all rounds (2002–2018). Similar patterns were observed cross-nationally and overtime for almost all cases as higher levels of the overall measurements of openness to change, self-transcendence and self-enhancement were obtained for those inclined to the higher end of satisfaction with life and happiness scales, respectively and to the higher end of general health scale. The reverse was true for the Spanish datasets of 2008 and 2018 for openness to change on satisfaction with life, the datasets of 2008 and 2018 for self-transcendence on happiness and the dataset of 2010 for self-enhancement on happiness.

This empirical study has certainly both strengths and limitations. The detailed demonstration of the proposed methodology for the case of unidimensional constructs and the required decisions for performing CFA based on current theory and practice should be noted among the strengths of the study. Initially, the items were considered as pseudo-interval and the criterion of corrected item-total correlations as recommended in the literature was used to decide on the inclusion of items in the analysis. In performing CFA, the appropriate method for model estimation was applied, multiple goodness-of-fit indices for model evaluation were used and specification searches were performed. Cronbach’s alpha and Composite reliability coefficients and convergent validity were estimated to assess the psychometric properties of the overall measurements as recommended in the literature.

Despite its strengths, the following limitations should be considered in drawing conclusions from this study. Firstly, as items were considered as pseudo-interval, the criterion of corrected item-total correlation was appropriately used, although the application of the same criterion on ordinal items too is not precluded by the literature (Chin et al. 2015; Hellström et al. 2019). Secondly, since the case of unidimensional constructs was under consideration, predetermined structures by theory were tested and CFA was employed for the full samples using the appropriate methods for pseudo-interval items. However, the polychoric matrix of associations should be employed in the case of ordinal items in future research (Brown 2015).

Thirdly, the resulting structures of four common items were similar but not quite as in a number of cases correlations between error variances differed and this should be dealt with in future research. Fourth, although Cronbach's alpha is widely used, it is considered as misestimator of scale reliability (Brown 2015). Therefore, composite reliability was used as an additional reliability coefficient whereas McDonald's Omega could be used instead (Zieger and Hagemann 2015). Fifth, although we demonstrated how the reliable and valid overall measurements maybe compared cross-nationally and overtime, their invariance should be tested before any such application. Sixth, many analytical decisions depended on the software used as pointed out in previous work (Michalopoulou 2017) and researchers may consider using different software in the future. Though SPSS is the most widely used software, for instance, in testing large samples for normality they provide only the poor-power Kolmogorov test (D'Agostino et al. 1990). In this respect, normality was assessed based on skewness and kurtosis where other software provides more sophisticated tests such as the D'Agostino-Pearson normality test (D'Agostino et al. 1990). Also, since SPSS Amos requires complete datasets, SPSS provides routines for the listwise and pairwise deletion of missing values, and mean substitution, regression imputation and Expectation maximization methods for their estimation. In the application, regression imputation was adopted which, as mentioned in previous work (Michalopoulou 2017) is problematic and Direct maximum likelihood should be considered as it is the best method for dealing with missing data under certain assumptions (Brown 2015; Enders 2010; Tabachnick and Fidell 2007). Finally, although by theory lower values indicate higher levels of openness to change, self-transcendence and self-enhancement, in future research it is best to reverse the order of scales' scores so as low and high values would indicate lower and higher levels, respectively.

The empirical methodology presented may be easily employed to other constructs defined as unidimensional by theory. As mentioned before, we demonstrated the proposed methodology by considering four unidimensional measurements so as to better discuss how it may be extended to cover the case of constructs defined by theory as multi-dimensional. In this respect, initially, their structure should be investigated and their psychometric properties assessed at country level and year as presented in great detail in our previous work (Charalampi 2018; Charalampi et al. 2019, 2020; Michalopoulou 2017). Then, the resulting valid and reliable subscales may be treated separately as unidimensional constructs and the sequence of decisions presented for creating the overall measurements may be easily applied bearing in mind that at least four items are required to define each measurement since "factors that are represented by two or three indicators may be underdetermined [...] and highly unstable across replications" (Brown 2015:21). However, it should be noted that this methodology is appropriate only for testing

models with first-order factors and therefore it rules out any considerations for models with second-order factors.

FUNDING

This research conducted under the auspices of the National Centre for Social Research is co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme “Human Resources Development, Education and Lifelong Learning 2014-2020” in the context of the project “Greece and Southern Europe: Investigating political trust to institutions, social trust and human values, 2002-2017” (MIS 5049524).

REFERENCES

- Bartholomew, David J., Fiona Steele, Irini Moustaki, and Jane Galbraith. 2008. *Analysis of multivariate social science data*. Chapman & Hall/CRC. <https://doi.org/10.1201/b15114>
- Bollen, Kenneth A. 1989. *Structural equations with latent variables*. Wiley. <https://doi.org/10.1002/9781118619179>
- Brown, Timothy A. 2015. *Confirmatory factor analysis for applied research* (2nd edition). The Guilford Press.
- Carey, Siobhan (Ed.) 2000. *Measuring adult literacy: The International Adult Literacy Survey (IALS) in the European context*. Office for National Statistics.
- Chang, Sea-Jin, Arjen van Witteloostuijn, and Lorraine Eden. 2010. “From the Editors: Common method variance in international business research.” *Journal of International Business Studies* 41(2): 178–184. <https://doi.org/10.1057/jibs.2009.88>
- Charalampi, Anastasia. 2018. *The importance of items’ level of measurement in investigating the structure and assessing the psychometric properties of multidimensional constructs* (Doctoral dissertation). Panteion University of Social and Political Sciences.
- Charalampi, Anastasia, Catherine Michalopoulou, and Clive Richardson. 2019. “Determining the structure and assessing the psychometric properties of multidimensional scales constructed from ordinal and pseudo-interval items.” *Communications in Statistics – Case Studies, Data Analysis and Applications* 5(1): 26–38. <https://doi.org/10.1080/23737484.2019.1579683>
- Charalampi, Anastasia, Catherine Michalopoulou, and Clive Richardson. 2020. “Validation of the 2012 European Social Survey measurement of wellbeing in seventeen European countries.” *Applied Research in Quality of Life* 15(1): 73–105. <https://doi.org/10.1007/s11482-018-9666-4>
- Chin, Weng Yee, Edmond P. Choi, Kit T. Y. Chan, and Carlos K. H. Wong 2015. “The psychometric properties of the Center for Epidemiologic Studies depression scale in Chinese primary care patients: Factor structure, construct validity, reliability, sensitivity and responsiveness.” *PLoS ONE* 10(8): e0135131. <https://doi.org/10.1371/journal.pone.0135131>

- Cieciuch, Jan, and Eldad Davidov. 2012. "A comparison of the invariance properties of the PVQ-40 and the PVQ-21 to measure human values across German and Polish samples." *Survey Research Methods* 6(1): 37–48.
- Clark, Lee A., and David Watson. 1995. "Constructing validity: Basic issues in objective scale development." *Psychological Assessment*, 7(3): 309–319. <https://doi.org/10.1037/1040-3590.7.3.309>
- D'Agostino, Ralph B., Albert Belanger and Ralph B. Jr. D'Agostino, 1990. "A suggestion for using powerful and informative tests of normality." *The American Statistician* 44(4): 316–321. <https://doi.org/10.1080/00031305.1990.10475751>
- Davidov, Eldad. 2008. "A cross-country and cross-time comparison of the human values measurements in the second Round of the European Social Survey." *Survey Research Methods* 2(1): 33–46.
- Davidov, Eldad, Peter Schmidt, and Shalom H. Schwartz. 2008. "Bringing values back in the adequacy of the European Social Survey to measure values in 20 countries." *Public Opinion Quarterly* 72(3): 420–445. <https://doi.org/10.1093/poq/nfn035>
- Ehling, Manfred. 2003. "Harmonising data in official statistics." Pp. 17–31 in *Advances in cross-national comparison: A European working book for demographic and socio-economic variables*, edited by J. H. P. Hoffmeyer-Zlotnik and C. Wolf. Kluger Academic-Plenum. https://doi.org/10.1007/978-1-4419-9186-7_2
- Enders, Craig K. 2010. *Applied Missing Data Analysis*. The Guilford Press.
- Ertanir, Beyhan, Christian Rietz, Ulrike Graf, and Wassilis Kassis. 2021. "A cross-national validation of the shortened version of the Adolescent Stress Questionnaire (ASQ-S) among adolescents from Switzerland, Germany and Greece." *Frontiers in Psychology* 12: 1–11. <https://doi.org/10.3389/fpsyg.2021.619493>
- European Social Survey. n.d. The human values scale: Findings from the European Social Survey. Retrieved from: http://www.europeansocialsurvey.org/docs/findings/ESS_Findings_HVS.pdf
- European Social Survey Round 1 Data. 2002. Data File Edition 6.6. Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
- European Social Survey Round 2 Data. 2004. Data File Edition 3.6. Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
- European Social Survey Round 4 Data. 2008. Data File Edition 4.5. Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
- European Social Survey Round 5 Data. 2010. Data File Edition 3.4. Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
- European Social Survey Round 6 Data. 2012. Data File Edition 2.4. Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
- European Social Survey Round 8 Data. 2016. Data File Edition 2.1. Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
- European Social Survey Round 9 Data. 2018. Data File Edition 2.0. Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.

- Fornell, Claes, and David F. Larcker. 1981. "Evaluating structural equation models with unobservable variables and measurement error." *Journal of Marketing Research* 18(1): 39–50. <https://doi.org/10.1177/002224378101800104>
- Harkness, Janet A., Ana Villar, and Brad Edwards. 2010. "Translation, adaptation and design." Pp. 117–140 in *Survey methods in multinational, multiregional, and multicultural contexts*, edited by J.A. Harkness, M. Braun, B. Edwards, T.P. Johnson, L.P. Lyberg, P.Ph. Mohler, B.E. Pennell, and T.W. Smith (Eds.). Wiley. <https://doi.org/10.1002/9780470609927.ch7>
- Hellström, Amanda, Peter Hagell, Anders Broström, Martin Ulander, Annemarie I. Luik, Colin A. Espie, and Kristofer Årestedt. 2019. "A classical test theory evaluation of the sleep condition indicator accounting for the ordinal nature of item response data." *PLoS ONE* 14(3): e0213533. <https://doi.org/10.1371/journal.pone.0213533>
- Hu, Li-tze, and Peter M. Bentler. 1999. "Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives." *Structural Equation Modeling* 6(1): 1–55. <https://doi.org/10.1080/10705519909540118>
- Kish, Leslie. 1994. "Multi-population survey designs: Five types with seven shared aspects." *International Statistical Review* 62(2): 167–186. <https://doi.org/10.2307/1403507>
- Lilleoja, Laur, and Willem E. Saris. 2014. "Testing a new operationalization of the basic values on Estonian-and Russian-speaking subpopulations in Estonia." *Social Indicators Research* 116(1): 153–172. <https://doi.org/10.1007/s11205-013-0272-4>
- Lilleoja, Laur, and Willem E. Saris. 2015. "Does correction for measurement error have an effect on the structure and comparability of basic human values?" *Survey Research Methods* 9(3): 169–187.
- Marsh, Herb W., Kit-Tai Hau, and Zhonglin Wen. 2004. "In search of golden rules: Comment on hypotheses-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings." *Structural Equation Modeling* 11(3): 320–341. https://doi.org/10.1207/s15328007sem1103_2
- Michalopoulou, Catherine. 2016. "Statistical internationalism: From Quetelet's census uniformity to Kish's cross-national sample survey comparability." *Statistical Journal of the IAOS* 32(4): 545–554. <https://doi.org/10.3233/SJI-160960>
- Michalopoulou, Catherine. 2017. "Likert scales require validation before application – Another cautionary tale." *Bulletin de Méthodologie Sociologique* 134: 5–23. <https://doi.org/10.1177/0759106317693786>
- Michalopoulou, Catherine, and Maria Symeonaki. 2017. "Improving Likert scale raw scores interpretability with k-means clustering." *Bulletin de Méthodologie Sociologique* 135: 101–109. <https://doi.org/10.1177/0759106317710863>
- Nunnally, Jum C., and Ira H. Bernstein. 1994. *Psychometric theory*. McGraw-Hill.
- Raykov, Tenko. 2007. "Reliability if deleted, not 'alpha if deleted': Evaluation of scale reliability following component deletion." *British Journal of Mathematical and Statistical Psychology* 60(2): 201–216. <https://doi.org/10.1348/000711006X115954>
- Sagin, Lilach, and Shalom H. Schwartz. 2000. "Value priorities and subjective well-being: Direct relations and congruity effects." *European Journal of Social Psychology* 30(2): 177–198. [https://doi.org/10.1002/\(SICI\)1099-0992\(200003/04\)30:2<177::AID-EJSP982>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1099-0992(200003/04)30:2<177::AID-EJSP982>3.0.CO;2-Z)

- Schmitt, Thomas A. 2011. "Current methodological considerations in exploratory and confirmatory factor analysis." *Journal of Psychoeducational Assessment* 29(4): 304–322. <https://doi.org/10.1177/0734282911406653>
- Schwartz, Shalom H. n.d. "Computing scores for the 10 human values." Retrieved from https://www.europeansocialsurvey.org/docs/methodology/ESS_computing_human_values_scale.pdf
- Schwartz, Shalom H. 1992. "Universals in the content and structure of values: Theory and empirical tests in 20 countries." Pp. 1–65 in *Advances in experimental social psychology*, edited by M. Zanna. Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60281-6](https://doi.org/10.1016/S0065-2601(08)60281-6)
- Schwartz, Shalom H. 2012. *An overview of the Schwartz theory of basic values*. Online readings in psychology and culture, Unit 2. Retrieved from <http://scholarworks.gvsu.edu/orpc/vol2/iss1/11>. <https://doi.org/10.9707/2307-0919.1116>
- Tabachnick, Barbara G., and Linda S. Fidell. 2007. *Using multivariate statistics*. Pearson Allyn & Bacon.
- The ESS Sampling Expert Panel. 2016. *Sampling guidelines: Principles and implementation for the European Social Survey*. ESS ERIC Headquarters. Retrieved from <http://www.europeansocialsurvey.org/>
- Thompson, Bruce. 2005. *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. American Psychological Association. <https://doi.org/10.1037/10694-000>
- West, Stephen G., John F. Finch, and Patrick J. Curran. 1995. "Structural equation models with nonnormal variables: Problems and remedies." Pp. 56–75 in *Structural equation modeling: Concepts, issues, and applications*, edited by R. H. Hoyle. Sage.
- Zieger, Matthias, and Dirk Hagemann. 2015. "Editorial: Testing the unidimensionality of items." *European Journal of Psychological Assessment* 31(4): 231–237. <https://doi.org/10.1027/1015-5759/a000309>

APPENDIX

Table A1: The demographic and social characteristics of Southern Europeans aged 18+ in the European Social Survey (ESS) 2002–2018 as compared to the IPUMS-International microdata of the 2002 and 2011 censuses

Country	<i>N</i>	Men (%)	Women (%)	Age (s)	Married (%)	Secondary education or lower (%)	In paid work* (%)
Greece							
IPUMS							
2001	8,509,646	48.6	51.4	47.0 (18.26)	63.5	79.9	46.3
2011	8,926,161	48.2	51.8	49.2 (18.66)	60.9	76.9	42.2
ESS							
2002	2,511	43.1	56.9	50.4 (18.77)	63.1	85.7	41.7
2004	2,363	43.6	56.4	50.7 (18.58)	62.5	81.1	39.0
2008	2,019	45.1	54.9	45.8 (16.29)	59.5	73.5	58.1
2010	2,649	43.5	56.5	48.4 (18.34)	55.8	75.2	39.7
Italy							
IPUMS							
2001	47,161,075	47.7	52.3	49.0 (17.95)	61.1	92.0	45.5
2011	50,396,628	47.8	52.2	49.5 (17.23)	55.4	88.5	46.5
ESS							
2004	1,497	49.2	50.8	48.6 (17.62)	55.6	87.0	46.3
2012	935	48.6	51.4	47.9 (17.79)	54.2	77.0	50.8
2016	2,524	48.7	51.3	50.1 (18.30)	53.7	86.4	49.2
2018	2,685	47.2	52.8	52.1 (18.93)	52.7	82.0	44.5
Portugal							
IPUMS							
2001	8,451,180	47.6	52.4	46.5 (18.52)	66.4	91.9	55.3
2011	8,657,240	47.0	53.0	49.1 (18.53)	55.9	84.6	50.4
ESS							
2002	1,444	41.1	58.9	49.3 (18.12)	62.3	88.9	51.8
2004	2,008	39.5	60.5	50.1 (19.00)	57.6	89.2	43.3
2008	2,296	38.6	61.4	53.9 (19.19)	56.9	87.4	41.6
2010	2,104	39.9	60.1	54.9 (18.59)	56.8	88.5	38.0
2012	2,117	39.3	60.7	57.9 (18.60)	50.4	89.2	39.8
2016	1,249	41.6	58.4	52.7 (17.85)	87.4	72.7	48.3
2018	1,034	41.9	58.1	53.1 (17.74)	51.4	70.1	49.1

Spain

IPUMS

2001	33,505,967	48.5	51.5	46.2 (18.65)	57.1	86.1	44.6
2011	38,458,315	48.4	51.6	50.4 (18.60)	57.5	82.7	43.4

ESS

2002	1,670	47.1	52.9	49.8 (18.58)	58.9	77.6	43.6
2004	1,608	51.1	48.9	46.1 (18.24)	60.1	72.0	54.9
2008	2,486	47.5	52.5	47.9 (18.60)	56.9	76.0	54.1
2010	1,840	49.2	50.8	46.6 (17.88)	53.0	70.0	49.2
2012	1,841	48.3	51.7	48.4 (17.53)	56.0	72.2	44.9
2016	1,918	49.6	50.4	50.3 (17.77)	55.7	66.7	49.7
2018	1,613	50.3	49.7	49.6 (17.98)	51.0	66.2	51.4

* The reference period for main activity was defined as during the last 7 days.

Table A2: Harman's single-factor test results for Southern Europe: European Social Survey, 2002–2018

Country	% of variance explained						
	2002	2004	2008	2010	2012	2016	2018
Greece	25.959	28.297	28.645	23.123	-	-	-
Italy	-	21.275	22.515	28.318	29.622	-	-
Portugal	28.693	33.646	30.008	33.816	30.599	23.585	24.947
Spain	26.712	23.671	20.637	18.097	20.146	18.969	20.358