Measurement Invariance of the SOC-13 Sense of Coherence Scale Across Gender and Age Groups

Grevenstein, Dennis; Bluemke, Matthias

Multistudy Report

Measurement Invariance of the SOC-13 Sense of Coherence Scale Across Gender and Age Groups

Dennis Grevenstein[1] and Matthias Bluemke[2]

[1]Psychological Institute, Heidelberg University, Germany
[2]GESIS – Leibniz Institute for the Social Sciences, Mannheim, Germany

Abstract: Sense of coherence (SOC) describes an individual's ability to deal with life challenges (manageability), comprehend the environment (comprehensibility), and perceive life and its challenges as meaningful (meaningfulness). We examine measurement invariance (MI) of the SOC-13 scale across gender and age groups in a matched sample of N = 1,816 (50% females; age range 16–83 years). A two-factor model, with a common factor for manageability/comprehensibility items and a second factor for meaningfulness items, best represented the SOC-13 in all groups. Full metric, partial scalar, and full strict invariance held across gender groups. Across age groups, full metric, partial scalar, and partial strict invariance could be established. We conclude that SOC-13 is a reliable and valid measure. Measurement is comparable across gender and age.

Keywords: sense of coherence, factorial validity, measurement invariance

OPEN MATERIALS

Sense of coherence (SOC) represents the core concept in Antonovsky's salutogenic theory (Antonovsky, 1979, 1987). Antonovsky proposed that health and disease should not be considered binary on/off states. Instead, he assumed that every human being can be placed on a larger continuum between health and disease, and SOC represents the most crucial concept that helps people move to the health end of the continuum. He argued that stressors are so ubiquitous in life that humans need a range of *general resistance resources* to deal with stressors and life challenges (Antonovsky, 1987). Thus, salutogenesis offers a resource-oriented perspective on health.

SOC comprises three components:

(1) manageability describes an individual's feeling that one has the necessary behavioral capacity and resources (e.g., skills, family, a social network) to deal with life challenges;
(2) comprehensibility is a cognitive aspect that represents an individual's perception that internal aspects and external situations and events are rational and understandable, and that even chaotic situations can be structured; and
(3) meaningfulness reflects that life has some kind of (emotional) meaning, so that its demands and challenges are worthy of investment and engagement (Eriksson & Mittelmark, 2017).

Antonovsky considered SOC an "orientation to life," rather than a temperamental personality trait (Antonovsky, 1987). He theoretically explained the development of SOC as a dynamic process up to age 30. Up to this age, SOC is supposed to be fluctuant, malleable, and shaped by experience. Consistency (enhancing comprehensibility), load-balancing (enhancing manageability), and participation in decision-making (enhancing meaningfulness) are all supposed to foster SOC across the developmental phase (Antonovsky, 1987). This theoretical view was backed up by empirical research according to a recent review of 37 studies from 14 countries, in which the authors concluded that "[t]he ... surveyed studies support the conceptualization of the SOC construct as an important personal resource that develops during childhood" (Idan et al., 2017, p. 118). Adolescence is seen as a particularly sensitive phase. Notably, the quality of parent-child relationships (Rivera et al., 2013) and a child-centered parenting style (Feldt et al., 2005) have been shown to be main predictors of SOC. SOC was also correlated with the quality of family relationships in later life (Grevenstein et al., 2019).

https://doi.org/10.1027/1015-5759/a000641

The clinical utility of SOC has been shown many times. SOC has been linked to good mental health and health-related behavior (Eriksson & Lindström, 2006), general psychological well-being (Nilsson et al., 2010), depression (Haukkala et al., 2013), anxiety (Moksnes, Espnes, & Haugan, 2013), general psychological distress (Grevenstein, Aguilar-Raab, et al., 2016), burnout (Grevenstein et al., 2018), satisfaction with life (Moksnes, Løhre, & Espnes, 2013), and substance use of tobacco, alcohol, and cannabis (Grevenstein, Bluemke, et al., 2016).

Historically, SOC has been criticized for its alleged similarity to classic personality traits like the Big Five personality factors, specifically neuroticism or emotional stability (Geyer, 1997). High negative correlations with neuroticism have been found, as well as smaller positive correlations with extraversion, agreeableness, and conscientiousness (Feldt, Metsäpelto, et al., 2007; Hochwälder, 2012; Kase et al., 2018). Overall, the Big Five can predict up to 40% of the SOC variance (Hochwälder, 2012). SOC has also shown surprisingly high longitudinal stability, almost comparable to temperamental personality traits. For adults, test-retest reliabilities of .78 over 1 year, .59–.67 over 5 years, and .54 over 10 years emerged (Eriksson & Lindström, 2005). Even in adolescence, SOC was found to be moderately stable. Honkinen and colleagues (2008) found only minor changes of mean SOC scores between ages 15 and 18 longitudinally. SOC at age 15 also predicted SOC scores at age 24 longitudinally at $\beta = .59$ on a latent level (Grevenstein & Bluemke, 2017). Nonetheless, SOC has a unique value for the prediction of health outcomes. SOC has shown incremental validity for the prediction of health-related outcomes above and beyond the Big Five traits (Grevenstein et al., 2018; Grevenstein & Bluemke, 2015), dispositional optimism, resilience, self-compassion (Grevenstein, Aguilar-Raab, et al., 2016), and mindfulness (Grevenstein et al., 2018).

Taken together, current results indicate that changes in SOC cannot be simply traced back to a sensitive period at a specific age. Feldt and colleagues compared two groups of 25- to 29-years-old and 35- to 40-years-old participants regarding their longitudinal change in SOC (Feldt et al., 2003). Both groups improved in a similar fashion with the older group showing very slightly lower SOC. In the much larger Finnish HeSSup study, participants over the age of 30 showed consistently higher mean SOC scores than participants under the age of 30 (Feldt, Lintula, et al., 2007). Silverstein and Heap (2015) showed that mean SOC scores increased continuously with age for older adults beyond the age of 55. Yet all this work presupposed (and left untested) the belief that the SOC scores can legitimately be compared across age groups. It is a rather strong assumption to think that the interpretation of SOC items and the applicability of SOC scale throughout ontogenesis are possible without measurement bias.

The SOC-13 Scale

The most popular measure of SOC is the 13-item SOC scale, originally published by Antonovsky (1987). The scale has been validated in a range of later studies (Antonovsky, 1993) and is widely accepted as a reliable and valid measure of SOC (Eriksson & Lindström, 2005). Still, the factorial validity of the SOC-13 scale has also been debated extensively in the past. Antonovsky developed the scale at a time before the general availability of software packages for structural equation modeling (SEM) and confirmatory factor analysis (CFA). Nonetheless, several studies have demonstrated the theoretically derived three-factor structure of the SOC scale (Feldt et al., 2000, 2003), though difficulties emerged. In many cases, items of the SOC scale had to be dropped or the measurement model had to be modified (Feldt et al., 2000, 2003).

One modification to the factor model has consistently and substantially improved model fit (Feldt et al., 2003; Moksnes & Haugan, 2014; Richardson et al., 2007): A residual correlation between items #2 and #3 (#2 "... were you surprised by the behavior of people who you thought you knew well?"; #3: "... have people you counted on disappointed you?") has been interpreted to reflect an additional aspect of interpersonal trust shared between the items (Frenz et al., 1993). Only recently have several studies replicated the intended three-factor model with all items included while allowing for one pair of correlated residuals between items 2 and 3 (Grevenstein, Aguilar-Raab, et al., 2016; Moksnes & Haugan, 2014; Stern et al., 2019). In all cases the comprehensibility and manageability factors correlated so highly ($r > .94$) that one can question if they are truly conceptually distinct.

As an alternative, a more parsimonious two-factor model has been proposed with one factor spanning comprehensibility and manageability items, and with meaningfulness constituting the second factor, though not to be mistaken as a secondary factor of lesser importance (Grevenstein, Aguilar-Raab, et al., 2016; Grevenstein & Bluemke, 2017; Zimprich et al., 2006). Despite its parsimony, this two-factor model has also shown superior predictive validity (Grevenstein, Aguilar-Raab, et al., 2016). One explanation that suggests itself is that comprehensibility and manageability are reciprocal aspects of conquering life stressors, yet meaningfulness constitutes a distinct, but an equally important component that provides the motivation to mobilize any coping resources (Antonovsky, 1987). This is in line with research showing that having a feeling of purpose in life helps people rise above mental health issues (Park, 2010).

Measurement Invariance

A comparison of SOC scores across situational contexts, measurement times, or groups of participants requires that the measurement model is valid across the different subsamples (Vandenberg & Lance, 2000). One needs to ascertain that manifest SOC scores reflect the same latent construct to the same degree. Stability indices and comparisons of manifest mean scores are only valid across different groups as far as strict measurement invariance (MI) can be established (Chen, 2008). Observed differences in scale means have to reflect true differences in latent means, not different item utilization or item difficulties. Otherwise, latent group differences or associations between a latent variable and external criteria might be explained by dissimilar measurement.

MI of the SOC scale has been addressed in the past, with some limitations. Comparing samples of Caucasian Americans and Asian Americans, Stein and colleagues (2006) had to drop items to achieve an acceptable model fit. Across two age groups of adolescents (12–14 years and 15–18 years), Zimprich et al. (2006) showed strict MI of a two-factor model, yet they retained all the items of the SOC-13 scale. Feldt and colleagues showed longitudinal MI across a 5-year span but tested only the invariance of factor loadings and item residuals (Feldt, Lintula, et al., 2007). Grevenstein and Bluemke (2017) showed longitudinal MI at age 15 and age 24, with partial scalar and strict invariance for a two-factor model that included all items. Luyckx and colleagues (2012), again, dropped two items from the SOC-13 scale when investigating MI across age and gender. Results supported scalar MI across ages 14–30 years. The authors also declared scalar invariance across gender, yet described a drop in model fit that – when following common heuristics for MI strictly – exceeded accepted cut-offs. Unfortunately, no detailed analyses of partial MI were provided, so model misfit could not be attributed to specific items. Hittner (2007) tested MI across gender when applying a single-factor model. Though all items were retained, only configural and metric MI were established, and scalar or strict MI were not even tested. To summarize, the field seems to be in a state of disarray, hence the need for a systematic investigation of MI of the SOC-13 scale across age and gender with the modified two-factor model that has consistently emerged in previous work.

Methods

Procedure and Participants

We pooled several samples collected between 2014 and 2016 to analyze the data presented in this study. In all studies, data were collected in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Though studies carried out at the Psychological Institute of the University of Heidelberg were granted an exemption from having to be run past the ethical review board at the time, unless research grant proposals were about to be submitted or sensitive topics were involved, informed consent had been obtained from all participants, and participation was completely voluntary. Most participants had completed short online studies without compensation.

We aimed to select a suitable sample for analysis from a total pool of N = 4,154 (78.1% female) German participants spanning a wide age range of 13–83 years. We used propensity score matching and the *R*-package “MatchIt” with its “nearest” algorithm (Ho et al., 2007) to select a sample with balanced gender groups matched for age. We initially had to exclude 13 participants from the pool, because they had not disclosed their gender. The final sample included N = 1,816 participants (50.0% female). Sample characteristics are displayed in Table 1. As intended, men and women did not differ with regard to their age, $t(1,814) = 0.09$, $p = .84$. Based on theoretical grounds we divided the sample into three age groups: Youth and young adults (age = 16–29 years; n = 1,008); adults (age = 30–49 years; n = 484), and older adults including seniors (age > 50 years; n = 324). There were no missing SOC-13 values, mostly due to the fact that the online participants were technically required to provide answers to every item; otherwise, they were considered drop-out participants, because we had assured them that refusing to participate (any further) was possible at any time.

Measures

We used the 13-item version of Antonovsky’s original Orientation to Life scale (Antonovsky, 1987). The German adaptation was provided by Schumacher and colleagues (2000). The scale includes five comprehensibility items (e.g., “Has it happened in the past that you were surprised by the behavior of people whom you thought you knew well?”), four manageability items (e.g., “Has it happened that people whom you counted on disappointed you?”), and four meaningfulness items (e.g., “Do you have the feeling that you do not really care about what goes on around you?”). Answers were provided on 7-point rating scales, marked from 1 = *very often* to 7 = *very seldom or never* most of the time. Items #1, #2, #3, and #7 were recoded before computing mean scores.

Statistical Analysis

We used SPSS 22 for descriptives and Mplus 7.4 (Muthén & Muthén, 1998–2012) for the CFAs. The arbitrariness of

Table 1. Descriptives, reliability, and standardized factor loadings for accepted models in the total sample (CFA), and subgroups gender (MGCFA 5) and age (MGCFA 5a)

	Total	Female	Male	Age 16–29	Age 30–49	Age 50–83
	N = 1,816	*n* = 908	*n* = 908	*n* = 1,008	*n* = 484	*n* = 324
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Age	33.48 (14.29)	33.45 (14.22)	33.51 (14.38)	22.83 (2.98)	38.87 (5.95)	58.55 (6.30)
SOC mean	4.75 (0.98)	4.73 (1.00)	4.77 (0.96)	4.51 (0.95)	4.87 (0.96)	5.32 (0.82)
SOC1	5.51 (1.45)	5.77 (1.35)	5.24 (1.51)	5.22 (1.52)	5.73 (1.33)	6.09 (1.18)
SOC2	4.02 (1.59)	3.88 (1.59)	4.15 (1.58)	4.07 (1.60)	3.95 (1.57)	3.94 (1.59)
SOC3	3.90 (1.69)	3.67 (1.67)	4.13 (1.68)	3.86 (1.73)	3.87 (1.70)	4.06 (1.54)
SOC4	5.25 (1.37)	5.40 (1.31)	5.11 (1.42)	5.16 (1.38)	5.17 (1.42)	5.68 (1.19)
SOC5	4.82 (1.62)	4.83 (1.61)	4.82 (1.64)	4.64 (1.62)	4.83 (1.67)	5.38 (1.45)
SOC6	4.94 (1.56)	4.96 (1.53)	4.93 (1.59)	4.57 (1.57)	5.17 (1.53)	5.76 (1.16)
SOC7	5.05 (1.25)	5.11 (1.23)	5.00 (1.26)	4.82 (1.22)	5.18 (1.25)	5.58 (1.13)
SOC8	4.67 (1.71)	4.53 (1.75)	4.81 (1.65)	4.18 (1.69)	5.05 (1.59)	5.65 (1.34)
SOC9	4.38 (1.82)	4.22 (1.86)	4.54 (1.76)	4.03 (1.80)	4.56 (1.81)	5.18 (1.58)
SOC10	4.43 (1.68)	4.36 (1.71)	4.50 (1.65)	4.25 (1.68)	4.40 (1.67)	5.03 (1.57)
SOC11	4.69 (1.48)	4.69 (1.52)	4.69 (1.44)	4.44 (1.51)	4.87 (1.43)	5.23 (1.28)
SOC12	4.91 (1.68)	5.03 (1.65)	4.78 (1.71)	4.52 (1.68)	5.15 (1.65)	5.74 (1.35)
SOC13	5.20 (1.60)	5.09 (1.67)	5.31 (1.52)	4.89 (1.67)	5.42 (1.49)	5.82 (1.26)
α	.87	.87	.86	.85	.87	.85
	λ	λ	λ	λ	λ	λ
SOC1	.35	.35	.35	.29	.34	.31
SOC2	.33	.33	.33	.36	.36	.33
SOC3	.47	.48	.48	.49	.49	.46
SOC4	.56	.56	.56	.55	.55	.50
SOC5	.57	.57	.57	.56	.56	.52
SOC6	.66	.65	.65	.61	.61	.71
SOC7	.73	.72	.72	.72	.72	.67
SOC8	.75	.75	.75	.71	.71	.76
SOC9	.77	.77	.77	.75	.75	.72
SOC10	.73	.73	.73	.74	.74	.70
SOC11	.49	.49	.49	.47	.47	.44
SOC12	.80	.80	.80	.79	.79	.75
SOC13	.69	.69	.69	.66	.66	.70
CR	.89	.89	.89	.88	.88	.87

Note. CR = Raykov's composite reliability.

using cut-offs notwithstanding, we evaluated model fit by (1) the – ideally nonsignificant – χ^2-test (Bentler & Bonett, 1980); (2) the comparative fit index (CFI) with values of .90/.95 and above indicating appropriate/good model fit (Bentler, 1990; Hu & Bentler, 1999); (3) the root mean square error of approximation (RMSEA) with values of .00–.05/.06–.08/.09–.10 indicating excellent/adequate/poor model fit (Browne & Cudeck, 1993); and (4) the standardized root mean square residual (SRMR) with values less than .08 (Hu & Bentler, 1999) or .05 (Schumacker & Lomax, 2010) considered to reflect good or excellent fit.

When comparing different models based on the same data and variables, we prefer the Akaike information criterion (AIC) and the Bayesian Information Criterion (BIC), which can be used to compare the quality of different models. Lower scores indicate better model fit (Akaike, 1987), and differences larger than 10 indicate "very strong" differences (Raftery, 1995). AIC commonly emphasizes accuracy, whereas BIC provides the best trade-off between accuracy and parsimony, which is most relevant for MI testing procedures. In line with prior research, we used Robust Maximum Likelihood (MLR) for parameter estimation.

For an estimation of reliability, we provide – for descriptive purposes – Cronbach's α, but also Raykov's composite reliability (CR; Raykov, 1997) as an SEM-based reliability estimate. As the SOC-13 scale lacks essential tau-equivalence and strict unidimensionality, Cronbach's α will be biased, whereas composite reliability is unbiased and preferable (Graham, 2006).

MI can be tested via a series of nested, increasingly restricted confirmatory factor-analytical (CFA) models (Meredith, 1993; Vandenberg & Lance, 2000). MI across independent groups, such as age or gender, is investigated using multiple-groups CFA (Brown, 2006). If age trends were assessed within the same group of participants, a longitudinal design for testing MI is required (Marsh & Grayson, 1994; Millsap & Cham, 2012). Four increasingly restrictive forms of MI can be tested (Vandenberg & Lance, 2000): (1) Configural MI indicates equal construct dimensionality and item-to-factor patterns across groups. Factor loadings, item intercepts, and residuals may differ. (2) Metric MI requires all factor loadings to be equal across groups. (3) Scalar MI additionally constrains all item intercepts to be equal across groups. (4) Strict MI further assumes equal residual variances. If at least some levels of MI can be established, it is possible to further investigate the invariance of structural parameters. Assuming metric invariance held, we wanted to test (5) the invariance of factor variances and covariances, including the residual correlation between items #2 and #3. As the last step, (6) the equality of factor means can be tested, assuming scalar invariance held.

Different levels of MI have ramifications for the applicability of scales and the comparability of scores. Metric MI indicates that latent variables representing the substantive factor reflect the same psychological meaning. Technically, metric MI implies that item scores are based on the same unit of measurement, which allows for a comparison of (latent) variance/covariance structures. Scalar MI denotes that item difficulties are comparable, which allows for a comparison of (latent) means. Strict MI indicates an equal impact of sources of item specificity (e.g., unreliability). When strict MI holds, differences in manifest variables are due to true differences in the latent variables, rather than item-specific measurement error. If strict MI holds, direct comparisons of manifest scale means are possible. In combination with equal latent variances, strict MI also implies that measurement reliability (proportion of true score variance to total variance) is comparable.

Tests of MI have often shown that invariance levels beyond metric invariance are hard to achieve (Schmitt & Kuljanin, 2008), but even lower levels of invariance may support comparable measurement (Tran, 2009). Traditionally, partial MI can be investigated if some item parameters (either loadings or intercepts) are non-invariant. For example, partial metric MI does not require all, but two, of the factor loadings to be equal (one anchor item's plus another invariant item's loadings). Partial scalar MI (one anchor item's plus another invariant item's intercepts) is statistically sufficient to compare latent means (Byrne et al., 1989; Lubke & Dolan, 2003). In a review on MI testing, Schmitt and Kuljanin (2008) concluded that "partial invariance made little difference in the estimates of structural model parameters." Candidate items that supply invariant model parameters can be found on the basis of χ^2-based Modification Indices (ModInd; Byrne et al., 1989). Researchers are advised to relax parameters one at a time to see if model fit can be improved and if partial MI can be established. A ModInd around 3.84 ($df = 1$) is just statistically significant at an arbitrary level of $p = .05$ for Type-I errors. Researchers are advised to look for modifications that substantially exceed this threshold (Brown, 2006). We generally followed these recommendations, and all modifications were executed based on ModInd. ModInd were used when a specific invariance test failed, but partial invariance was still an option. In an iterative manner, the largest ModInd (typically >10) was used to identify the model parameter most in need of being freed from a cross-group equivalence constraint. The partial invariance model was then inspected for acceptable model fit.

In MI testing, the alternative models which are nested in less constrained baseline models are compared based on χ^2-difference tests. MLR uses scaled χ^2-scores, but χ^2-difference scores are not χ^2-distributed themselves, necessitating Satorra-Bentler scaled χ^2-difference tests (Satorra, 2000; Satorra & Bentler, 2001). Any χ^2-tests or χ^2-difference tests are likely to reach significance due to our large sample. Independent from the sample size, model fit indices can be used to evaluate MI analyses. Going from one step to the next, a drop in CFI less or equal to .010 is conventionally considered acceptable unless there is a concurrent increase of RMSEA greater than +.015 (Chen, 2007; Cheung & Rensvold, 2002). However, strictly adhering to cut-offs when examining ΔCFI and ΔRMSEA is problematic when the simulation parameters used for deriving the cut-offs somehow differ from the present case (Fan & Sivo, 2009; Saris et al., 2009). A better alternative is to look for lower BIC values that indicate a better tradeoff between accuracy and parsimony, irrespective of sample size and model complexity.

Results

Descriptive Data Analysis

Men and women hardly differed at all regarding their mean SOC scores, $t(1{,}814) = 0.765$, $p = .45$, $d = 0.04$. Across age groups, SOC scores increased almost linearly with age, $F(2, 1{,}813) = 98.41$, $p < .001$, $\eta^2_p = .10$. With regard to reliability estimates, Cronbach's αs were consistently high in all groups. For comparison with a representative German sample, we computed a mean score of SOC means for 808 participants in the two older age groups ($M_{age} = 46.76$). SOC means were comparable to the reference sample (N = 2,005, M_{age} = 50.03) reported by Schumacher and colleagues (2000): M = 5.05 (SD = 0.93) versus M_{ref} = 5.01 (SD_{ref} = 0.89), d = 0.04.

Confirmatory Factor Analyses and Measurement Invariance

Matching previous researchers' considerations, we first compared a three-factor model and a two-factor model (across the whole sample), assuming a pair of correlated residuals between items #2 and #3. The three-factor model fitted the data well, $\chi^2(61)$ = 376.22, p < .001, RMSEA = .053, CI_{90} = [.048, .059], CFI = .952, SRMR = .036, but so did the two factor-model, $\chi^2(63)$ = 383.77, p < .001, RMSEA = .053, CI_{90} [.048, .058], CFI = .951, SRMR = .036. The factors for manageability and comprehensibility were statistically nearly indistinguishable (r = .96). Having replicated the recently emerged standard model for the SOC-13 scale, we accepted the two-factor model as the basis for the following MI analyses.

Results and model fit indices of the MGCFAs are depicted in Table 2. Testing first the invariance across genders, the initial configural invariance model showed a good model fit. Constraining factor loadings to be equal (metric MI) did not impair model fit. We observed a noticeable drop in model fit when testing scalar MI, so we investigated modification indices (ModInd) to check for potential adjustments of the model. The intercept of item #1 (ModInd = 37.80, $\Delta\chi^2$ = 38.47) "Do you have feeling that you do not really care about what goes on around you?" was not invariant across genders. After relaxing its equality constraint, partial scalar MI held. Within the context of a partial scalar invariant model, we next examined strict invariance by constraining residuals to be equal across groups. The decrease in model fit was well within acceptable levels. At the level of structural parameters, enforcing equal variances and covariances also did not harm model fit. At last, we tested for equal factor means. The drop in CFI was still below .010 and RMSEA increased only slightly. Still, ModInd indicated unequal factor means for both factors (meaningfulness: ModInd: 43.69; comprehensibility/manageability: ModInd: 41.34). On the basis of the variance invariance model (M5), we can quantify the estimated latent mean differences. Compared to females, males had lower (unstandardized) latent means on meaningfulness (Est. = −0.193, SE = 0.054, p < .001, d = .12) and higher latent means on comprehensibility/manageability (Est. = 0.161, SE = 0.051, p = .002, d = .10).

We then investigated MI across age groups. Again, the configural MI model showed good model fit. Metric invariance also held unconditionally. The scalar MI model showed a substantial decrease in model fit. ModInd indicated unequal intercepts for several items among the young adults: item #8 (ModInd = 49.29, $\Delta\chi^2$ = 49.71) "Do you have very mixed-up feelings and ideas?", item #6 (ModInd = 33.36, $\Delta\chi^2$ = 33.58) "Do you have the feeling that you are in an unfamiliar situation and do not know what to do?", item #1 (ModInd = 31.63, $\Delta\chi^2$ = 31.10), item #4 (ModInd = 16.98, $\Delta\chi^2$ = 8.41) "Until now your life has had ... clear goals", item #10 (ModInd = 16.02, $\Delta\chi^2$ = 16.85) "Many people – even those with a strong character – sometimes feel like sad sacks (losers) in certain situations. How often have you felt this way in the past?", item #2 (ModInd = 11.17, $\Delta\chi^2$ = 11.98) "Has it happened in the past that you were surprised by the behavior of people whom you thought you knew well?", and item #3 (ModInd = 25.17, $\Delta\chi^2$ = 26.09) "Has it happened that people whom you counted on disappointed you?". After these modifications, partial scalar MI could be established. Notably, all modifications pertained to the young adult group. We next tested strict MI by constraining all residuals to be equal across age groups, yet once more model fit dropped below accepted cut-offs. In the senior age group, the residuals of item #6 (ModInd = 31.81, $\Delta\chi^2$ = 42.93), item #8 (ModInd = 17.92, $\Delta\chi^2$ = 20.67), and item #13 (ModInd = 14.63, $\Delta\chi^2$ = 15.02) were non-invariant, as was item #1 in the young adult group (ModInd = 25.85, $\Delta\chi^2$ = 26.28). After modifications partial strict MI could be established.

At the level of structural parameters, we constrained all variances and covariances to be equal across groups. Model fit dropped slightly. Most notably, SRMR increased by .025. ModInd indicated unequal variances for both SOC factors in the senior group: meaningfulness (ModInd = 16.82, $\Delta\chi^2$ = 169.42) and comprehensibility/manageability (ModInd = 2,009.67, $\Delta\chi^2$ = 1029.24). Finally, we constrained latent means to be equal. As expected, model fit clearly decreased. On the basis of the accepted partial variance invariance model, we could estimate latent mean differences. Compared to the young adult group, the adult age group had higher (unstandardized) scores on latent meaningfulness (Est. = 0.297, SE = 0.071, p < .001, d = .19) and comprehensibility/manageability (Est. = 0.410, SE = 0.065, p < .001, d = .29) factors. Differences between the young and senior groups were even stronger for meaningfulness (Est. = 0.799, SE = 0.075, p < .001, d = .59) and comprehensibility/manageability (Est. = 0.902, SE = 0.068, p < .001, d = .74).

Discussion

The present research investigated measurement invariance (MI) of SOC as measured by the SOC-13 scale across age and gender. In line with prior research, a three-factor model

Table 2. Measurement invariance of SOC across gender and age

MGCFA comparison	Equal loadings	Equal intercepts	Equal residuals	Equal variances	Equal means	*df*	χ^2	Δdf	$\Delta\chi^2$	CFI	RMSEA [CI_{90}]	SRMR	AIC	BIC
Gender														
1. Configural invariance						126	432.58**	–	–	.954	.052 [.046, .057]	.037	80,385	80,837
2. Metric invariance	×					137	464.51**	11	31.27**	.951	.051 [.046, .056]	.045	80,396	80,787
3. Scalar invariance	×	×				148	560.16**	11	104.82**	.938	.055 [.051, .060]	.053	80,480	80,810
3a. Partial scalar invariance	×	×				147	521.69**	10	61.04**	.944	.053 [.048, .058]	.051	80,438	80,773
4. Strict invariance	×	×	×			160	563.62**	13	41.78**	.939	.053 [.048, .057]	.059	80,458	80,722
5. Structural: (co)variances	×	×	×	×		**164**	**575.09****	**4**	**11.44***	**.938**	**.053 [.048, .057]**	**.063**	**80,463**	**80,706**
6. Structural: means	×	×	×	×	×	166	629.58**	2	62.77**	.930	.055 [.051, .060]	.072	80,521	80,753
Age														
1. Configural invariance						189	530.97**	–	–	.944	.055 [.049, .060]	.042	79,969	80,646
2. Metric invariance	×					211	565.13**	22	35.37*	.942	.053 [.044, .058]	.052	79,968	80,524
3. Scalar invariance	×	×				233	787.46**	22	239.97**	.909	.058 [.058, .068]	.066	80,176	80,611
3a. Partial scalar invariance	×	×				226	609.73**	15	45.21**	.937	.053 [.048, .058]	.055	79,985	80,458
4. Strict invariance	×	×	×			252	781.58**	26	159.45**	.913	.059 [.054, .064]	.100	80,147	80,478
4a. Partial strict invariance	×	×	×			248	676.69**	22	66.47**	.930	.053 [.049, .058]	.070	86,294	86,599
5. Structural: (co)variances	×	×	×	×		256	709.98**	8	31.34**	.925	.054 [.049, .059]	.095	80,052	80,361
5a. Partial (co)variances	×	×	×	×		**254**	**683.99****	**6**	**8.17**	**.929**	**.053 [.048, .058]**	**.078**	**80,023**	**80,342**
6. Structural: means	×	×	×	×	×	258	843.24**	4	160.87**	.904	.061 [.057, .066]	.136	80,201	80,499

Note. Model fit: *$p < .05$; **$p < .001$. Accepted models printed in bold. χ^2 values are Satorra-Bentler scaled.

fitted the data, yet resulted in a very high correlation between the manageability and comprehensibility factors. For reasons of parsimony, we tested MI on the basis of a two-factor model (Grevenstein, Aguilar-Raab, et al., 2016; Grevenstein & Bluemke, 2017; Zimprich et al., 2006). The alleged correlated residuals between items #2 and #3 replicated across all subgroups (Feldt et al., 2003; Grevenstein, Aguilar-Raab, et al., 2016; Moksnes, Espnes, et al., 2013).

Across gender groups, full metric MI, partial scalar MI, and strict MI could be established. Even at the level of structural parameters, the invariance of all variances and covariances could be shown. Taken together, only one intercept (item #1) differed, and females endorsed the item more readily. Only minor differences in latent means emerged. Even manifest scale means appear to be quite comparable across genders without introducing a large amount of bias.

Across age groups, full metric MI, partial scalar MI, and partial strict MI held. Non-invariance at the scalar level was due to items in the young adult group, with seven items having non-invariant intercepts. This finding is unprecedented, but highly relevant, as not a single study has investigated SOC's MI across this wide age range before. This finding is highly illuminating for salutogenic theory. Even though non-invariance may be due to uniform item bias, or disparate use of items by young cohort members, a more theoretical explanation pertains to the development of SOC at young adult age.

Antonovsky described the development of SOC as a dynamic process up to the age of 30 (Antonovsky, 1987). He assumed SOC to be fully developed in later years. It has long been accepted that people face different developmental tasks across various age stages (Havighurst, 1972). Full metric MI for the SOC-13 provided the first evidence that the same psychological construct is measured irrespective of the age of participants. And yet, the different components underlying SOC (at the factor and item level) are unevenly important at various age stages. One can easily fathom that most life challenges are different for a 20-year-old who just started life on their own from a 50-year-old who has dealt with these challenges and may then be tightly embedded in family structures or professional settings. This logic is in line with recent research on life goals. Much like the challenges that life poses, life goals change across the life-span. The importance of personal-growth, status, and work goals decreased, whereas prosocial-engagement increased in importance (Bühler et al., 2019). Moreover, goal-adjustment, that is, adaptive disengagement and re-engagement capacities were found to predict individuals' well-being and health (Barlow et al., 2019). Our results of partial scalar MI have to be seen through the lens of developmental tasks that influence how people interpret the SOC-13 items from within their age-dependent context, which systematically affects the item difficulty when choosing one of the response options of the non-invariant items. This age-dependent differential item functioning needs to be kept in mind when comparing SOC scores, before arriving at firm conclusions on true differences between groups, or changes across the life-span, that involve young adult age.

Unfortunately, neither could we establish strict MI unconditionally, though only a minority of item residuals had to be estimated freely. Covariances between factors and between the residuals of items #2 and #3 were found to be equal, yet factor variances in the senior group were reduced. The latent means of both meaningfulness and comprehensibility/manageability showed a substantial, and nearly linear, increase along with age. This is in line with prior research, where older participants reported higher SOC means (Feldt, Lintula, et al., 2007; Silverstein & Heap, 2015). This increase of SOC scores refutes Antonovsky's age stability hypothesis, which entails variability and development of generalized resistance resources up to the age of 30, but stability afterward. This finding can easily be reconciled with previous research demonstrating relatively high stability of participants' rank-order. The increase in SOC that may occur in later life affects all people to a similar extent. Consequently, in the future researchers need to be clear about whether they refer to absolute SOC levels (which progress even at higher age) or relative individual differences (which appear quite stable). If one reads Antonovsky to refer to the former, the theoretical supposition has to be rejected, whereas it may be compatible with the latter notion.

Regarding the measurement of SOC, our results shine a positive light on the factorial validity of the SOC-13 scale. For both types of group comparisons, we could establish (partial) strict MI. Most researchers consider partial scalar MI as "good enough" for an interpretation of mean structures (Byrne et al., 1989; Schmitt & Kuljanin, 2008; Steenkamp & Baumgartner, 1998). Strict MI is even more difficult to achieve, yet the consequences of unequal residuals may be low in practical terms (Lubke & Dolan, 2003; Schmitt & Kuljanin, 2008).

Antonovsky conceived SOC as a meta-construct. As a medical-sociologist, he envisioned numerous potential external factors that could influence SOC (Antonovsky, 1987). Antonovsky assumed that a person's SOC develops in line with a broad range of concepts, including socioeconomic status, social environment, general intelligence, and personal experiences in interpersonal relationships. Our results support the conclusion that despite the complexity of the SOC construct, its measurement across age and gender is – almost surprisingly – comparable.

Limitations

Our study offers progress beyond prior research, as we were able to analyze a large sample with a wide range of age represented in it. Due to the use of propensity score matching, we can confidently assume that the effects of participants' age and gender were not confounded. Still, our sample does not represent a real probability-based sample. Also, it needs to be seen if good results can be achieved in other languages as well.

Conclusions

We confirmed partial strict MI of the SOC-13 scale across gender and age groups, allowing for comparable measurement of SOC. Our results lend support to previous and future comparisons of variance/covariance structures (correlations) and mean structures across groups when based on the SOC-13 scale.

References

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika, 52*(3), 317–332. https://doi.org/10.1007/BF02294359

Antonovsky, A. (1979). *Health, stress, and coping*. Jossey-Bass.

Antonovsky, A. (1987). *Unraveling the mystery of health: How people manage stress and stay well*. Jossey-Bass.

Antonovsky, A. (1993). The structure and properties of the Sense of Coherence scale. *Social Science & Medicine, 36*(6), 725–733. https://doi.org/10.1016/0277-9536(93)90033-Z

Barlow, M. A., Wrosch, C., & McGrath, J. J. (2020). Goal adjustment capacities and quality of life: A meta-analytic review. *Journal of Personality, 88*(2), 307–323. https://doi.org/10.1111/jopy.12492

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588–606. https://doi.org/10.1037/0033-2909.88.3.588

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. Guilford Press.

Browne, 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). Sage.

Bühler, J. L., Weidmann, R., Nikitin, J., & Grob, A. (2019). A closer look at life goals across adulthood: Applying a developmental perspective to content, dynamics, and outcomes of goal importance and goal attainability. *European Journal of Personality, 33*(3), 359–384. https://doi.org/10.1002/per.2194

Byrne, B. M., Shavelson, R. J., & Muthén, B. O. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*(3), 456–466. https://doi.org/10.1037/0033-2909.105.3.456

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology, 95*(5), 1005–1018. https://doi.org/10.1037/a0013193

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Eriksson, M., & Lindström, B. (2005). Validity of Antonovsky's Sense Of Coherence scale: A systematic review. *Journal of Epidemiology and Community Health, 59*(6), 460–466. https://doi.org/10.1136/jech.2003.018085

Eriksson, M., & Lindström, B. (2006). Antonovsky's Sense of Coherence Scale and the relation with health: A systematic review. *Journal of Epidemiology and Community Health, 60*(5), 376–381. https://doi.org/10.1136/jech.2005.041616

Eriksson, M., & Mittelmark, M. B. (2017). The sense of coherence and its measurement. In M. B. Mittelmark, S. Sagy, M. Eriksson, G. F. Bauer, J. M. Pelikan, B. Lindström, & G. A. Espnes (Eds.), *The handbook of salutogenesis* (pp. 97–106). Springer International Publishing.

Fan, X., & Sivo, S. A. (2009). Using Δgoodness-of-fit indexes in assessing mean structure invariance. *Structural Equation Modeling, 16*(1), 54–69. https://doi.org/10.1080/10705510802561311

Feldt, T., Kokko, K., Kinnunen, U., & Pulkkinen, L. (2005). The role of family background, school success, and career orientation in the development of sense of coherence. *European Psychologist, 10*(4), 298–308. https://doi.org/10.1027/1016-9040.10.4.298

Feldt, T., Leskinen, E., Kinnunen, U., & Mauno, S. (2000). Longitudinal factor analysis models in the assessment of the stability of sense of coherence. *Personality and Individual Differences, 28*(2), 239–257. https://doi.org/10.1016/s0191-8869(99)00094-x

Feldt, T., Leskinen, E., Kinnunen, U., & Ruoppila, I. (2003). The stability of sense of coherence: Comparing two age groups in a 5-year follow-up study. *Personality and Individual Differences, 35*(5), 1151–1165. https://doi.org/10.1016/s0191-8869(02)00325-2

Feldt, T., Lintula, H., Suominen, S., Koskenvuo, M., Vahtera, J., & Kivimäki, M. (2007). Structural validity and temporal stability of the 13-item Sense of Coherence Scale: Prospective evidence from the population-based HeSSup study. *Quality of Life Research, 16*(3), 483–493. https://doi.org/10.1007/s11136-006-9130-z

Feldt, T., Metsäpelto, R., Kinnunen, U., & Pulkkinen, L. (2007). Sense of coherence and five-factor approach to personality: Conceptual relationships. *European Psychologist, 12*(3), 165–172. https://doi.org/10.1027/1016-9040.12.3.165

Frenz, A. W., Carey, M. P., & Jorgensen, R. S. (1993). Psychometric evaluation of Antonovsky's Sense of Coherence Scale. *Psychological Assessment, 5*(2), 145–153. https://doi.org/10.1037/1040-3590.5.2.145

Geyer, S. (1997). Some conceptual considerations on the sense of coherence. *Social Science & Medicine, 44*(12), 1771–1779. https://doi.org/10.1016/S0277-9536(96)00286-9

Graham, J. M. (2006). Congeneric and (essentially) tau-equivalent estimates of score reliability: What they are and how to use them. *Educational and Psychological Measurement, 66*(6), 930–944. https://doi.org/10.1177/0013164406288165

Grevenstein, D., Aguilar-Raab, C., & Bluemke, M. (2018). Mindful and resilient? Incremental validity of sense of coherence over mindfulness and Big Five personality factors for quality of life outcomes. *Journal of Happiness Studies, 19*, 1883–1902. https://doi.org/10.1007/s10902-017-9901-y

Grevenstein, D., Aguilar-Raab, C., Schweitzer, J., & Bluemke, M. (2016). Through the tunnel, to the light: Why sense of coherence covers and exceeds resilience, optimism, and self-compassion.

Personality and Individual Differences, 98, 208–217. https://doi.org/10.1016/j.paid.2016.04.001

Grevenstein, D., & Bluemke, M. (2015). Can the Big Five explain the criterion validity of sense of coherence for mental health, life satisfaction, and personal distress? *Personality and Individual Differences, 77*, 106–111. https://doi.org/10.1016/j.paid.2014.12.053

Grevenstein, D., & Bluemke, M. (2017). Longitudinal factor analysis and measurement invariance of sense of coherence and general self-efficacy in adolescence. *European Journal of Psychological Assessment, 33*, 377–387. https://doi.org/10.1027/1015-5759/a000294

Grevenstein, D., Bluemke, M., & Kroeninger-Jungaberle, H. (2016). Incremental validity of sense of coherence, neuroticism, extraversion, and general self-efficacy: Longitudinal prediction of substance use frequency and mental health. *Health and Quality of Life Outcomes, 14*, Article 9. https://doi.org/10.1186/s12955-016-0412-z

Grevenstein, D., Bluemke, M., Schweitzer, J., & Aguilar-Raab, C. (2019). Better family relationships – higher well-being: The connection between relationship quality and health related resources. *Mental Health & Prevention, 14*, Article 200160. https://doi.org/10.1016/j.mph.2019.200160

Grevenstein, D. (2021). *SOC MI materials. European Journal of Psychological Assessment*. Zenodo. https://doi.org/105281/zenodo.4744808

Haukkala, A., Konttinen, H., Lehto, E., Uutela, A., Kawachi, I., & Laatikainen, T. (2013). Sense of coherence, depressive symptoms, cardiovascular diseases, and all-cause mortality. *Psychosomatic Medicine, 75*(4), 429–435. https://doi.org/10.1097/PSY.0b013e31828c3fa4

Havighurst, R. J. (1972). *Developmental tasks and education*. McKay.

Hittner, J. B. (2007). Factorial invariance of the 13-item Sense of Coherence Scale across gender. *Journal of Health Psychology, 12*(2), 273–280. https://doi.org/10.1177/1359105307074256

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis, 15*(3), 199–236. https://doi.org/10.1093/pan/mpl013

Hochwälder, J. (2012). The contribution of the Big Five personality factors to sense of coherence. *Personality and Individual Differences, 53*(5), 591–596. https://doi.org/10.1016/j.paid.2012.05.008

Honkinen, P., Suominen, S., Helenius, H., Aromaa, M., Rautava, P., Sourander, A., & Sillanpää, M. (2008). Stability of the sense of coherence in adolescence. *International Journal of Adolescent Medicine and Health, 20*(1), 85–91. https://doi.org/10.1515/ijamh.2008.20.1.85

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. https://doi.org/10.1080/10705519909540118

Idan, O., Braun-Lewensohn, O., Lindström, B., & Margalit, M. (2017). Salutogenesis: Sense of coherence in childhood and in families. In M. B. Mittelmark, S. Sagy, M. Eriksson, G. F. Bauer, J. M. Pelikan, B. Lindström, & G. A. Espnes (Eds.), *The Handbook of Salutogenesis* (pp. 107–121). Springer International Publishing.

Kase, T., Ueno, Y., & Oishi, K. (2018). The overlap of sense of coherence and the Big Five personality traits: A confirmatory study. *Health Psychology Open, 5*(2), 2055102918810654. https://doi.org/10.1177/2055102918810654

Lubke, G. H., & Dolan, C. V. (2003). Can unequal residual variances across groups mask differences in residual means in the common factor model? *Structural Equation Modeling, 10*(2), 175–192. https://doi.org/10.1207/S15328007SEM1002_1

Luyckx, K., Goossens, E., Apers, S., Rassart, J., Klimstra, T., Dezutter, J., & Moons, P. (2012). The 13-item Sense of Coherence Scale in Dutch-speaking adolescents and young adults: structural validity, age trends, and chronic disease. *Psychologica Belgica, 52*(4), 351–368. https://doi.org/10.5334/pb-52-4-368

Marsh, H. W., & Grayson, D. (1994). Longitudinal confirmatory factor analysis: Common, time-specific, item-specific, and residual-error components of variance. *Structural Equation Modeling: A Multidisciplinary Journal, 1*(2), 116–145. https://doi.org/10.1080/10705519409539968

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), 525–543. https://doi.org/10.1007/BF02294825

Millsap, R. E., & Cham, H. (2012). Investigating factorial invariance in longitudinal data. In B. Laursen, T. D. Little, & N. A. Card (Eds.), *Handbook of developmental research methods* (pp. 109–127). Guilford Press.

Moksnes, U. K., Espnes, G. A., & Haugan, G. (2013). Stress, sense of coherence and emotional symptoms in adolescents. *Psychology & Health, 29*(1), 32–49. https://doi.org/10.1080/08870446.2013.822868

Moksnes, U. K., & Haugan, G. (2014). Validation of the orientation to life questionnaire in norwegian adolescents, construct validity across samples. *Social Indicators Research, 119*(2), 1105–1120. https://doi.org/10.1007/s11205-013-0536-z

Moksnes, U. K., Løhre, A., & Espnes, G. A. (2013). The association between sense of coherence and life satisfaction in adolescents. *Quality of Life Research, 22*(6), 1331–1338. https://doi.org/10.1007/s11136-012-0249-9

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide: Statistical analysis with latent variables* (7th ed.). Muthén & Muthén.

Nilsson, K. W., Leppert, J., Simonsson, B., & Starrin, B. (2010). Sense of coherence and psychological well-being: Improvement with age. *Journal of Epidemiology and Community Health, 64*(4), 347–352. https://doi.org/10.1136/jech.2008.081174

Park, C. L. (2010). Making sense of the meaning literature: An integrative review of meaning making and its effects on adjustment to stressful life events. *Psychological Bulletin, 136*(2), 257–301. https://doi.org/10.1037/a0018301

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25*, 111–163. https://doi.org/10.2307/271063

Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement, 21*(2), 173–184. https://doi.org/10.1177/01466216970212006

Richardson, C. G., Ratner, P. A., & Zumbo, B. D. (2007). A test of the age-based measurement invariance and temporal stability of Antonovsky's Sense of Coherence Scale. *Educational and Psychological Measurement, 67*(4), 679–696. https://doi.org/10.1177/0013164406292089

Rivera, F., García-Moya, I., Moreno, C., & Ramos, P. (2013). Developmental contexts and sense of coherence in adolescence: A systematic review. *Journal of Health Psychology, 18*(6), 800–812. https://doi.org/10.1177/1359105312455077

Saris, W. E., Satorra, A., & van der Veld, W. M. (2009). Testing structural equation models or detection of misspecifications? *Structural Equation Modeling, 16*, 561–582. https://doi.org/10.1080/10705510903203433

Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. H. Heijmans, D. S. G. Pollock, & A. Satorra (Eds.), *Innovations in multivariate statistical analysis. A Festschrift for Heinz Neudecker* (pp. 233–247). Kluwer Academic.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507–514.

Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review, 18*(4), 210–222. https://doi.org/10.1016/j.hrmr.2008.03.003

Schumacher, J., Wilz, G., Gunzelmann, T., & Brähler, E. (2000). Die Sense of Coherence Scale von Antonovsky – Teststatistische Überprüfung in einer repräsentativen Bevölkerungsstichprobe und Konstruktion einer Kurzskala [The Antonovsky Sense of Coherence Scale. Test statistical evaluation in a representative population sample and construction of a brief scale]. *Psychotherapie, Psychosomatik, Medizinische Psychologie, 50*(12), 472–482. https://doi.org/10.1055/s-2000-9207

Schumacker, R. E., & Lomax, R. G. (2010). *A beginners guide to structural equation modeling*. Routledge.

Silverstein, M., & Heap, J. (2015). Sense of coherence changes with aging over the second half of life. *Advances in Life Course Research, 23*, 98–107. https://doi.org/10.1016/j.alcr.2014.12.002

Steenkamp, J. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25*(1), 78–90. https://doi.org/10.1086/209528

Stein, J. A., Lee, J. W., & Jones, P. S. (2006). Assessing cross-cultural differences through use of multiple-group invariance analyses. *Journal of Personality Assessment, 87*(3), 249–258. https://doi.org/10.1207/s15327752jpa8703_05

Stern, B., Socan, G., Rener-Sitar, K., Kukec, A., & Zaletel-Kragelj, L. (2019). Validation of the Slovenian version of short Sense of Coherence Questionnaire (SOC-13) in multiple sclerosis patients. *Slovenian Journal of Public Health, 58*(1), 31–39. https://doi.org/10.2478/sjph-2019-0004

Tran, T. V. (2009). *Developing cross-cultural measurement invariance*. Oxford University Press.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4–69. https://doi.org/10.1177/109442810031002

Zimprich, D., Allemand, M., & Hornung, R. (2006). Measurement invariance of the abridged Sense of Coherence Scale in adolescents. *European Journal of Psychological Assessment, 22*(4), 280–287. https://doi.org/10.1027/1015-5759.22.4.280

History
Received August 1, 2019
Revision received September 18, 2020
Accepted January 7, 2021
Published online May 26, 2021
EJPA Section / Category Positive Psychology & Assessment

Open Science
We report how we determined our sample size, all data exclusions (if any), all data inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all measures in the study, and all analyses including all tested models. If we use inferential tests, we report exact *p* values, effect sizes, and 95% confidence or credible intervals.
Open Data: The information needed to reproduce all of the reported results are not openly accessible.
Open Materials: I confirm that there is sufficient information for an independent researcher to reproduce all of the reported methodology (Grevenstein, 2021).
Preregistration of Studies and Analysis Plans: This study was not preregistered.

Funding
Open access publication enabled by Heidelberg University.

ORCID
Dennis Grevenstein
https://orcid.org/0000-0002-1422-0309

Dennis Grevenstein
Psychological Institute
Heidelberg University
Hauptstr. 47-51
69117 Heidelberg
Germany
dennis.grevenstein@psychologie.uni-heidelberg.de