Psychiatry Research 272 (2019) 491–498

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

Factor structure models of the SCL-90-R: Replicability across community samples of adolescents

Antonio Preti[a,b,c,*], Mauro Giovanni Carta[a], Donatella Rita Petretto[b]

[a] Center for Liaison Psychiatry and Psychosomatics, University Hospital, University of Cagliari, Cagliari, Italy
[b] Section on Clinical Psychology, Department of Education, Psychology, Philosophy, University of Cagliari, Cagliari, Italy
[c] Genneruxi Medical Center, Cagliari, Italy

ARTICLE INFO

Keywords:
Symptom checklist (SCL-90-R)
Confirmatory factor analysis
Bifactor model
Psychological distress
Adolescents

ABSTRACT

Derogatis' Symptom Checklist-90-Revised (SCL-90-R) is one of the most widely used measures of psychological distress in both the clinical and the research settings, although its factor structure is still debated. In this study, the factor structure of the SCL-90-R has been investigated in two independent samples of Italian adolescents. Two samples of 817 (urban sample) and 507 (rural sample) adolescents attending high schools were involved. Confirmatory factorial analysis (CFA) was used alongside hierarchical nested, progressively constrained models to assess configural, metric and scalar invariance of the best models fitted by CFA. The standard nine-factor structure of the SCL-90-R resulted reproducible and invariant between the two samples, in both its correlated and hierarchical second-order implementations. Estimated reliability of the nine scales of the SCL-90-R was optimal. This study also confirmed the reproducibility of the bifactor models of the SCL-90-R with nine orthogonally independent factors and with nine correlated primary factors, which have been tested in some recent studies. Overall, the SCL-90-R measures both common and unique features of psychological distress in community samples. The measurement invariance across different levels of psychological distress in the factor structure of the SCL-90-R is an issue deserving further testing and investigation.

1. Introduction

Derogatis' Symptom Checklist-90-Revised (SCL-90-R) is one of the most widely used measures of psychological distress in both the clinical and the research settings (Derogatis and Cleary, 1977; Derogatis, 1994). The SCL-90-R enquires about psychological distress, usually over the latest seven/fourteen days, across nine subscales (somatization, obsessive-compulsive behavior, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism), and other minor disturbances (e.g., sleep disturbances). Additional global indexes can be computed (Global Severity Index [GSI]; Positive Symptom Total [PST]; and Positive Symptom Distress Index [PSDI]), with the GSI being the best indicator of the current level of overall psychological distress (Prunas et al., 2012).

The SCL-90-R items are simple and describe - in an easy-to-understand manner - the occurrence of psychological (e.g., "Feeling easily annoyed or irritated"), cognitive (e.g., "Trouble remembering things"), or somatic symptoms (e.g., "Pains in heart or chest"). This makes the SCL-90-R less invasive than other questionnaires that directly enquire about symptoms of mental disorders, and is therefore more suitable for use in adolescents. Another benefit of the SCL-90-R is its ability in measuring both internalizing (e.g., depression, anxiety) and externalizing (e.g., hostility) symptoms, which makes the SCL-90-R a helpful tool to investigate these dimensions of psychopathology in adolescents (Achenbach et al., 2016).

1.1. Age of onset of mental disorders and the measurement of psychological distress

Most mental disorders have their onset in adolescence and early youth. About a half of lifelong mental disorders had their onset by the mid-teens, which rate rises to three-fourths by the mid-20s (Kessler et al., 2007). For the most severe disorders such as schizophrenia spectrum psychosis or bipolar disorder, prodromal symptoms typically precede the onset of the full-blown syndrome (Geoffroy and Scott, 2017; Nelson et al., 2017; McGorry et al., 2018). Most of these prodromal symptoms are subthreshold and their degree of severity does not allow reaching a specific diagnosis. A consequence is the long delay between the onset of the disorder and its management, particularly for schizophrenia spectrum psychosis (Penttilä et al., 2014), and bipolar

* Corresponding author at: Centro Medico Genneruxi, Via Costantinopoli 42, Cagliari 09129, Italy
E-mail address: apreti@tin.it (A. Preti).

https://doi.org/10.1016/j.psychres.2018.12.146
Received 9 July 2018; Received in revised form 1 December 2018; Accepted 28 December 2018
Available online 28 December 2018
0165-1781/

disorder (Dagani et al., 2017). There is a relationship between the delay in treatment start and the clinical and functional outcomes in schizophrenia spectrum psychosis (Santesteban-Echarri et al., 2017) and bipolar disorder (Drancourt et al., 2013). However, there is very limited evidence that the duration of untreated psychosis can be shortened by applying the controlled interventional studies tested thus far (Oliver et al., 2018). Regarding bipolar disorder the information is insufficient to determine whether the duration of the untreated bipolar disorder can be modified (Dagani et al., 2017).

The most recent conceptualizations of the early intervention model posit that the very early stage of the prodromal phases should be the target of interventions aimed at improving the course of severe and potentially lifelong mental disorders (McGorry and van Os, 2013; Raballo et al., 2017). A kaleidoscopic series of micro-phenotypes, transitioning from anxiety and depression to subtle abnormalities of salience and perception and intermixing with emotional dysregulation and impulsivity, is likely to precede the convergence of the symptoms into more definite macro-phenotypes, the precursors of the currently defined, prototypical mental disorders (McGorry et al., 2018). These micro-aggregates of subthreshold symptoms could be missed by questionnaires and interviews aimed at more structured symptoms, such as those currently used to screen people with high at-risk mental states (e.g., Schultze-Lutter et al., 2015; Savill et al., 2018). A self-report tool, able to cover a wide array of symptoms and being simple enough in its format to be easy-to-use, such as the SCL-90-R, might be more suited to detect the psychological distress that may evolve into more defined high at-risk mental states.

There is evidence that the psychological distress accompanying psychotic experiences - but not the psychotic experiences per se - is an important predictor of non-suicidal self-injury and of suicide attempts in adolescents (Martin et al., 2015). Generally there is evidence that the scores on the scales of the SCL-90-R that tap into symptoms of depression and psychosis are associated with an enhanced risk of common mental disorders (Rössler et al., 2011a; 2011b). It is still a matter of debate whether the scores on the SCL-90-R represent a unique latent trait of psychological distress, as supposed for other measures of mild symptoms of anxiety, depression and psychotic experiences (Stochl et al., 2015; see also Caspi and Moffitt, 2018), or whether they convey a multidimensional description of the candidates' psychopathology, as suggested by its creators (Derogatis and Cleary, 1977; Derogatis, 1994).

1.2. Factor structure of the SCL-90-R

The replicability of the nine-subscale factor structure of the SCL-90-R was highly debated in the past. A 1985 review based on exploratory factor analysis studies declared that the standard nine-factor structure of the SCL-90-R was irreproducible across samples (Cyr et al., 1985). Several studies based on confirmatory factor analysis (CFA) failed, too, in retrieving the standard nine-factor structure of the SCL-90-R. It was suggested that failure in retrieving the standard nine-factor structure of the SCL-90-R in CFA studies was probably related to inadequate sample size, or to the use of inappropriate methods to fit the data in the presence of severe deviation from multivariate normal distribution, or still to inadequacy in dealing with the ordinal nature of the responses on the items (Urbán et al., 2014). When more adequate and robust methods for parameter estimation were applied to samples – that were large enough to allow the convergence of the algorithms - the standard nine-factor structure of the SCL-90-R showed an acceptable level of fit (e.g., Urbán et al., 2014; Sereda and Dembitskyi, 2016; Arrindell et al., 2017). More recently, a bifactor implementation of the standard nine-factor structure of the SCL-90-R proved to have a good fit in adult (Urbán et al., 2014) and adolescent samples (Rytilä-Manninen et al., 2016) from the general population, in psychiatric samples accessing a psychotherapy treatment program (Urbán et al., 2016), and in a large sample ($n > 2500$) of patients diagnosed with substance use and impulse control disorders (Arrindell et al., 2017).

Bifactor models fall in the class of hierarchical factor models. In a classic hierarchical second-order factor model, the co-variances of the primary factors are explained by the higher-order factor(s). In a bifactor model, the general and the group factors are constrained to be orthogonal (Holzinger and Swineford, 1937; Schmid, 1957). The general factor reflects the common elements among the items, and represents the individual differences on the target dimension (Reise et al., 2010; Reise, 2012). The group factors mirror the item response variance that is not accounted for by the general factor, and represent additional variance that is explained by sub-dimensions within the items. An issue with the bifactor models is their propensity to overfitting, i.e. their tendency towards capturing unwanted noise in the data (Bonifay et al., 2016). For this reason several benchmark indicators were developed to check for reasonable unidimensionality of the general factor extracted from a bifactor model. Among these benchmarks, the most frequently used are the explained common variance (ECV), the percentage of uncontaminated correlations (PUC), and Omega Hierarchical (OH) (Rodriguez et al., 2016). ECV is the ratio of the variance explained by the general factor to the variance explained by the model (i.e., the variance explained by the general factor plus the variance explained by the group factors). The PUC is the ratio of the number of uncontaminated correlations to the number of unique correlations. The OH reflects the percentage of the systematic variance in unit-weighted total scores that can be attributed to the individual differences in the general factor. As a rule, ECV is expected to be higher than 0.70; however, with PUC > 0.80, general ECV values are less important in predicting bias (Rodriguez et al., 2016).

The reproducibility of the bifactor model implementation of the standard nine-factor structure of the SCL-90-R is uncertain. Urbán et al. (2016) were able to get a good fit of the bifactor model implementation of the standard nine-factor structure of the SCL-90-R in a Hungarian sample ($n = 972$), but they failed to reproduce it in a Dutch sample ($n = 1902$). In the Dutch sample the standard bifactor model with nine specified factors did not reach the threshold for good fit on the comparative fit index (CFI) and the Tucker-Lewis Index (TLI), and reached the threshold for acceptable fit (> 0.90) on the CFI but not on the TLI in a respecification of the model. The TLI incorporates a relatively substantial penalty for model complexity and is scarcely affected by sample size (Marsh et al., 2005), thus it was identified as particularly useful in testing the fit of a bifactor model (Gignac, 2016).

The reproducibility of the factor structure of the SCL-90-R across independent samples was poorly investigated in the past. Reproducibility of a factor model cannot merely rely on the fit of the model in independent samples. Measurement invariance should be tested, too. The assessment of measurement invariance serves the purpose of demonstrating that the participants, assigned to groups of interest (in this case, to independent samples), interpret the single queries and the underlying latent factor in the same way (Jöreskog, 1971; Vandenberg and Lance, 2000). We decided to investigate the factor structure of the SCL-90-R across two independent samples of Italian adolescents, one from an urban area and another from a rural area.

1.3. Urban-rural differences in psychological distress

Growing evidence points towards the presence of urban-rural differences in the prevalence of mental disorders (Peen et al., 2010; Vassos et al., 2016). There is some evidence than the urban environment, when defined according to objective measures, shows appreciable associations with psychological distress, and with mood and anxiety symptoms in particular (Gong et al., 2016). Poor neighborhood quality, limited availability of green spaces, intense industrial activity and traffic volumes, they were all positively related to psychological distress (Gong et al., 2016). Most of the studies have been made thus far in European or US samples and, indeed, more recent evidence have suggested that urban-rural differences do vary depending on the type of disorder or symptom, age, town, geographical scale, and investigated country. A large US study, which included 55,583 adolescents aged 12-17 years and 116,459 adults aged 18 and older from the National Survey of Drug Use and Health, did not find urban-rural differences in the prevalence of major depression in adolescents, but found only a slightly higher prevalence of major depression and serious mental illness in adults living in small urban or semi-rural areas, with no relevant differences between the more clearly characterized urban areas and the rural areas (Breslau et al., 2014). A study that included 18,000 adults aged 18 years and older from a survey made in nine republics of the former Soviet Union, found that people living in small towns or villages had higher odds for psychological distress than those living in large cities, an opposite trend to what was observed in Western studies (Stickley et al., 2015). In a study including 4,209 adults aged 50 years and residing in Ghana, and 3,148 adults aged 50 years living in South Africa, no significant urban-rural differences were found in the 12-month prevalence of self-reported depression (Adjaye-Gbewonyo et al., 2018).

Overall the observation of differences by geographical areas in the prevalence of psychological distress, or lack thereof, especially when measured with self-report tools, might depend on the fact that people from rural areas give a different meaning to the interview/questionnaire items than people from urban areas. Some studies made in a closely related field (satisfaction with life) provide evidence that people's understanding of the intended meaning of a question, or their interpretation of certain words hence the conceptualization of an entire construct, may vary due to cultural attributions (Veenhoven, 1996). Since even subtle differences may affect the equality of constructs across groups, measurement invariance needs to be checked before comparing mean scores across groups (Vandenberg and Lance, 2000; Bieda et al., 2017).

1.4. Aims of the study

In this study four structural models of the SCL-90-R were tested: a) the unidimensional model, which assumes that the items tap into a single factor that is expected to represent a common source of variance; b) the standard nine-factor structure of the SCL-90-R; c) the second-order model, in which a shared variance is deemed to be attributable to a dimension external to the correlated sub-domains; and d) the bifactor implementation of the standard nine-factor structure of the SCL-90-R, with both the orthogonal and oblique rotations of the nine factors. Measurement invariance was then applied to confirm that the models with the best fit in the two independent samples also showed configural, metric, and scalar invariance across participants, whether from urban or rural areas.

2. Methods

Data were collected from two independent samples, from an urban area (1,851 inhabitants/Km^2) and a rural area (82 inhabitants/Km^2). The urban area is a metropolitan town surrounded by a network of suburbs and small towns, with over 250,000 inhabitants at the time of the study; the rural area corresponds to a rural town with nearly 30,000 inhabitants at the time of the study, a smaller population than the 50,000 inhabitants suggested for inclusion in the "intermediate area" according to the Eurostat classification of the degree of urbanisation, and it does not lie next to a highly populated area ("suburban area") (Eurostat, 2018).

All studies took place in Sardinia, a large island of Italy in the middle of the Mediterranean Sea. The competent institutional review boards approved the study protocol in accordance with the guidelines of the 1995 Declaration of Helsinki and its revisions (World Medical Association, 2013).

2.1. Participants and procedures

The surveys involved high-school students. Schools were stratified in order to achieve a representative sample of young people in terms of social and cultural backgrounds. In Sardinia about 85% of all adolescents attend high school after compulsory school, thus the samples included in the study can be considered representative of their peers. The surveyed high schools were a humanistic school, a science school, a technical-administrative school, and a vocational school (hotel management). A random selection was applied to the schools operating in the urban area to make the sample compatible with the one from the rural area (which included all the types of schools on the territory). The three upper classes of each course (age 15 to 19) were invited to take part in the study. A fraction of the students did not agree to participate in the study, and/or were absent from school during the study. Attrition rate was 17.2% in the urban area and 11.3% in the rural one.

Students were individually asked to participate in the study after having been approached as a group in their classrooms. Several contacts were provided to the students, in case they felt that they had the problems that were described in the questionnaire. Confidentiality was guaranteed, and informed consent was obtained from each student.

Table 1 summarizes the main characteristics of the two samples.

2.2. Measures

The Italian version of the SCL-90-R was used in the study. Each item is rated on a five-point Likert scale of distress from 0 (not at all) to 4 (extremely). Items are grouped into nine scales; the scores of each item are summed up and the resulting score is averaged by the number of items in the scale, so that each scale is rated from 0 (not at all) to 4 (extremely). For clinical purposes, raw scores can be converted into T-score by referring to the group-specific population-based norm tables in the manual.

Socio-demographic variables were collected with a self-report schedule with fixed queries on gender, age, etc. As a measure of socioeconomic status we used the highest level of their parents' education attainment (Galobardes et al., 2006), which was further classified into three categories: lower than high school diploma, high school diploma, college graduate or higher.

2.3. Statistical analysis

All data were coded and analyzed using the Statistical Package for Social Sciences (SPSS) version 20. Additional analyses were carried out in R (Core Team, 2017), using dedicated packages. All tests were two-tailed, with alpha set at $p < 0.05$. Means with standard deviations were reported for continuous variables. Counts and percentages were

Table 1
Sociodemographic characteristics of the samples included in the study.

Sample size	Urban sample $N = 817$	Rural sample $N = 507$	Statistics
Sex			$\chi^2 = 6.45$; df = 1; $p = 0.011$
Girls	535 (65%)	296 (58%)	
Boys	282 (35%)	211 (42%)	
Age			
Mean (SD) Median	17.6 (1.4) 17	17.4 (1.3) 17	$t = 2.23$; df = 1322; $p = 0.026$
Parental education			$\chi^2 = 47.9$; df = 2; $p < 0.0001$
Compulsory school	280 (34.3%)	232 (45.8%)	
High school diploma	301 (36.8%)	210 (41.4%)	
College graduate or higher	236 (28.9%)	65 (12.8%)	

reported for categorical variables.

Reliability of the SCL-90-R was estimated from the confirmatory factor analysis (CFA). Estimated Cronbach's α and McDonald's omega were both reported. Despite harsh criticism of the usefulness of the Cronbach's α (Sijtsma, 2009), its reporting is still advised (see Mokkink et al., 2010); Cronbach's α was therefore reported to allow comparison with past studies.

CFA was carried out with the *lavaan* package (Rosseel, 2012) running in R (R Core Team, 2012). The *lavaan* package has been shown to generate the same results as other software packages (Narayanan, 2012).

Mardia's test (Mardia, 1970) revealed violation of multivariate normality in both the urban (skew = 296566.70, $p < 0.0001$; kurtosis = 318.23, $p < 0.0001$) and the rural samples (skew = 259320.30, $p < 0.0001$; kurtosis = 204.11, $p < 0.0001$). To deal with violations of multivariate normal distribution and the ordinal nature of the responses on the items, the diagonally weighted least squares (DWLS) estimator was used. The *DWLS* approach automatically uses the WLS estimator with *polychoric* correlations as an input to create the asymptotic covariance matrix. The DWLS estimator is known to have a fairly good control of Type I error rates with sample sizes as small as 200 (Flora and Curran, 2004). However, to distinguish a bifactor model from a second-order model, a sample size around 500 might be required (Chen et al., 2006).

As in past studies (Urbán et al., 2014; Urbán et al., 2016), the items that do not belong to the main nine scales of the SCL-90-R (e.g., those on sleep disturbances) were not included in the models. In both the bifactor and the second-order models the error terms associated with each item were uncorrelated. As in some past studies (Urbán et al., 2014; 2016), we have tested the standard bifactor model, with the general factor and the specific factors completely orthogonal to each other (i.e., uncorrelated), and a bifactor model with correlated specific factors, still maintaining that the specific factors were uncorrelated to the general factor (i.e., all correlations of the specific factors with the general factor were fixed to zero).

Parameters for fit estimation were: the chi-square, the comparative fit index (CFI), the TLI, the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR). McDonald's omega was also calculated according to the model (McDonald, 1978). RMSEA values of 0.08 or lower, SRMR values of 0.09 or lower, CFI values of 0.90 or higher, and TLI values of 0.95 or higher are considered acceptable (Browne and Cudeck, 1993; Hu and Bentler, 1999). McDonald's omega around 0.90 is considered acceptable.

In the bifactor model of the SCL-90-R with nine orthogonally independent factors, the PUC is 0.89. The ECV and the OH were calculated with a Microsoft Excel-based calculator developed by Dueber (2016). For OH, a threshold ≥ 0.20 was used to accept relative independence of the group factor from the general factor, as in Urbán et al. (2014) and in Urbán et al. (2016). This conservative threshold was preferred over the more stringent threshold ≥ 0.50 proposed by Reise et al. (2013) since, as suggested by Urbán et al. (2016), the number of specific factors and the test length are likely to impact on the values of the hierarchical omega.

Model identification was verified according to the Bekker et al. (1994) method. Essentially the model is identified when the rank of the Jacobian matrix is equal to the number of free parameters in the factorial model, i.e. the number of columns in the matrix (see details in Preti et al., 2015).

Measurement invariance by location (urban versus rural) of the best model(s) was explored according to Byrne and van de Vijver (2010) using the R *semTools* package (Pornprasertmanit et al., 2013). Typically, a hierarchical set of steps should be followed to test invariance, starting from the identification of a well-fitting baseline model and then establishing successive equivalence constraints in the model parameters across the groups. Configural, metric and scalar invariance was tested. Configural invariance refers to whether the same CFA model is valid in each group. Metric invariance concerns the equivalence of the factorial loadings across groups. Scalar invariance is assumed when the item intercepts and the factor loadings are equally constrained across groups. When intercepts invariance is confirmed, then the latent means in both groups can be compared. Models were compared on the basis of changes in CFI, TLI and RMSEA (delta-CFI, delta-TLI, and delta-RMSEA). To accept or reject measurement invariance the following criteria were used: the fit of the model, and the conventionally agreed threshold of 0.01 for delta-CFI and delta-TLI, and 0.02 for delta-RMSEA (Cheung and Rensvold, 2002; Chen, 2007).

3. Results

The urban sample included a modest excess of girls compared to the rural sample, and a modest difference in age was noted. Socio-economic status was higher in the urban sample than in the rural one (Table 1).

Levels of psychological distress, as measured by the GSI, did not differ between the urban and the rural samples: 0.84 ± 0.57 versus 0.82 ± 0.55 ($t = 0.80$; df = 1322; $p = 0.42$).

3.1. Confirmatory factor analysis in the urban and rural samples

Results of the CFA in the two samples are listed in Table 2.

All models showed good fit according to the predefined parameters.

In the urban sample the ECV and OH of the standard bifactor model were, respectively, 0.77 and 0.98, representing optimal acceptability of the general factor as a sum score. In the rural sample the ECV of the standard bifactor model was low (0.12), but the OH was high (0.95).

ECV and OH of the bifactor model with oblique rotation (i.e., with correlated specific factors) were 0.75 and 0.97 in the urban sample, and 0.74 and 0.97 in the rural sample, suggesting a better fit of the bifactor model with oblique rotation than the standard bifactor model in the rural sample.

In the urban sample the relative omega was higher than a conventional threshold of 0.20 for the somatization (0.31), hostility (0.43), and phobic anxiety (0.36) scales in the standard bifactor model. In the rural sample, too, the relative omega was higher than 0.20 for the somatization (0.31), hostility (0.42), and phobic anxiety (0.31) scales in the standard bifactor model. In the bifactor model with oblique rotation, the relative omega was higher than a conventional threshold of 0.20 for the interpersonal sensitivity (0.31), hostility (0.38) and the phobic anxiety (0.37) scales in the urban sample, and for the somatization (0.46) and the hostility (0.49) scales in the rural sample. In all bifactor models, the hostility scale would contribute least to the general factor's variance in both the urban and the rural samples.

In both samples, the estimated Cronbach's α was close to the estimated omega, with lower values yet still above the conventional threshold of 0.70 for the phobic anxiety and the paranoid ideation scales in the urban sample, and for the hostility, the phobic anxiety, and the paranoid ideation scales in the rural sample (see Tables A1 to A6 in the supplementary material).

Considering the correlated, standard nine-factor structure of the SCL-90-R and its second-order implementation, in both the urban and the rural samples most items had a loading > 0.500 on their a priori factors/scales, confirming a good fit of the models (Tables A1 and A2 in the supplementary material; the hierarchical second-order implementation of this model is not reproduced, since it is conceptually and numerically equivalent to the simple correlated nine-factor model).

As far as the bifactor models are concerned, the loading of the items on the general factor was excellent and always statistically significant, while the loading of the items on the specific factors were as expected the less statistically significant the lower their OH (details in Tables A3, A4, A5, and A6 in the supplementary material).

On the basis of the fit indexes, all models except the unidimensional one can be considered equivalent and were tested for measurement

Table 2
Goodness-of-fit indices of the tested models.

Urban sample	Goodness of fit indicators χ^2	df	*p*	CFI	TLI	RMSEA (90%CI)	SRMR	McDonald ω	Wald rank rule *n* columns	Rank	
One-factor model	4690.58	3320	0.0001	0.991	0.991	0.022 (0.021 – 0.024)	0.053	0.97	166	166	Identified
Nine first-order model	3793.92	3284	0.0001	0.997	0.997	0.014 (0.012 – 0.016)	0.047	0.97	202	202	Identified
Second-order factor model (with nine primary factors)	4037.29	3311	0.0001	0.995	0.995	0.016 (0.015 – 0.018)	0.049	0.97	175	175	Identified
Bifactor model with nine orthogonally independent factors	3655.41	3237	0.0001	0.997	0.997	0.013 (0.010 – 0.015)	0.046	0.97	249	249	Identified
Bifactor model with nine correlated primary factors	3115.87	3201	0.857	1.000	1.000	0.000 (0.000 – 0.004)	0.043	0.97	285	285	Identified
Rural sample	χ^2	df	*p*	CFI	TLI	RMSEA (90%CI)			*n* columns	Rank	
One-factor model	2963.76	3320	1.00	1.00	1.00	0.00 (0.00 – 0.00)	0.056	0.97	174	174	Identified
Nine first-order model	2481.37	3284	1.00	1.00	1.00	0.00 (0.00 – 0.00)	0.051	0.97	202	202	Identified
Second-order factor model (with nine primary factors)	2650.11	3311	1.00	1.00	1.00	0.00 (0.00 – 0.00)	0.053	0.97	175	175	Identified
Bifactor model with nine orthogonally independent factors	2442.59	3237	1.00	1.00	1.00	0.00 (0.00 – 0.00)	0.051	0.97	249	249	Identified
Bifactor model with nine correlated primary factors	2045.35	3201	1.00	1.00	1.00	0.00 (0.00- 0.00)	0.046	0.97	285	285	Identified
Threshold for good fit			> 0.05	> 0.90	> 0.95	< 0.08	< 0.09	> 0.90	*n* =	rank	

invariance between the two datasets (urban and rural samples).

3.2. Measurement invariance of the factor models of the SCL-90-R by location

Fit was acceptable from the configural to the scalar model in the standard correlated nine factors, the second-order factor model with nine primary factors, and in the bifactor model with nine orthogonally independent factors (Table 3).

The bifactor model with oblique rotation (i.e., with correlated specific factors) had an exceptional fit. This might depend on the fact that this model is the most congruent with the data, or the model suffers from Type II error (failure to reject the null hypothesis of no difference across the increasingly constrained models) because of high sample size requirement.

Across models, delta-CFI, delta-TLI and delta-RMSEA were within the threshold for acceptance of measurement invariance. Overall, the results support configural, metric and strong invariance by location of the tested factor models of the SCL-90-R.

4. Discussion

The main finding of this study is the demonstration that the standard nine-factor structure of the SCL-90-R is reproducible in both the correlated and the second-order implementations. This is important on a clinical ground, since it supports the use of the nine scales to profile the psychological status of the candidates, and of the global indexes as an overall measure of their psychological distress. The study also

Table 3
Fit indices for measurement invariance tests of the factor structure of the SCL-90-R.
Comparison by location: Urban area (n = 817) versus Rural area (n = 507).

Nine first-order model	χ^2	df	*p*	CFI	TLI	RMSEA (90%CI)	delta-CFI	delta-TLI	delta-RMSEA
Configural invariance	6275.29	6568	0.995	1.000	1.000	0.000 (0.000 – 0.000)			
Metric invariance	7623.74	6642	0.0001	0.996	0.996	0.015 (0.013 – 0.017)	0.004	0.004	0.015
Scalar invariance	7709.24	6716	0.0001	0.996	0.996	0.015 (0.013 – 0.017)	0.000	0.000	0.000
Second-order factor model (with nine primary factors)	χ^2	df	*p*	CFI	TLI	RMSEA (90%CI)	delta-CFI	delta-TLI	delta-RMSEA
Configural invariance	6687.40	6622	0.284	1.000	1.000	0.004 (0.000 – 0.008)			
Metric invariance	8285.45	6704	0.0001	0.993	0.993	0.019 (0.017 – 0.020)	0.006	0.006	0.015
Scalar invariance	8371.10	6777	0.0001	0.993	0.993	0.019 (0.017 – 0.020)	0.000	0.000	0.000
Bifactor model with nine orthogonally independent factors	χ^2	df	*p*	CFI	TLI	RMSEA (90%CI)	delta-CFI	delta-TLI	delta-RMSEA
Configural invariance	6097.83	6474	1.000	1.000	1.000	0.000 (0.000 – 0.000)			
Metric invariance	7726.19	6630	0.0001	0.995	0.995	0.016 (0.014 – 0.017)	0.005	0.005	0.016
Scalar invariance	7805.57	6703	0.0001	0.995	0.995	0.016 (0.014 – 0.017)	0.000	0.000	0.000
Bifactor model with nine correlated independent factors	χ^2	df	*p*	CFI	TLI	RMSEA (90%CI)	delta-CFI	delta-TLI	delta-RMSEA
Configural invariance	5085.90	6402	1.000	1.000	1.000	0.000 (0.000 – 0.000)			
Metric invariance	6177.60	6558	1.000	1.000	1.000	0.000 (0.000 – 0.000)	0.000	0.000	0.000
Scalar invariance	6237.9	6631	1.000	1.000	1.000	0.000 (0.000 – 0.000)	0.000	0.000	0.000
Threshold for good fit			> 0.05	> 0.90	> 0.95	< 0.08	< 0.01	< 0.01	< 0.02

confirms the good reliability of the nine scales of the SCL-90-R, a finding consistent with initial studies (Derogatis and Savitz, 1999) and with more recent validation studies (Tomioka et al., 2008; Sereda and Dembitskyi, 2016).

Despite the evidence, several authors still have doubts about the structural indeterminacy of the SCL-90-R (Prunas et al., 2012). This happens mostly as consequence of inadequate methods being applied to conduct confirmatory factor analysis in some past studies. For example Vassend and Skrondal (1999), in an oft-quoted study on the topic, used Maximum Likelihood (ML) estimation to fit the data, but ML is not appropriate for ordinal data. Vassend and Skrondal (1999) also re-estimated the models with the asymptotically distribution-free (ADF) method, which requires very large samples (N > 5000) to produce unbiased estimates and is outperformed by the DWLS estimator. When more adequate and robust methods than the ML estimator were applied to parameter estimations, the reproducibility of the standard nine-factor structure of the SCL-90-R was confirmed by CFA (Urbán et al., 2014; Sereda and Dembitskyi, 2016; Arrindell et al., 2017).

The factors and scales in the SCL-90-R have always been reported to be substantially highly intercorrelated (e.g., Arrindellet al., 2006). Most critics of the standard nine-factor structure of the SCL-90-R refer to the high intercorrelations among the different scales as a reason to discount the multidimensional factor structure of the tool (e.g., Kevin Chapman et al., 2012). Another commonly reported finding is the extraction - by principal component analysis - of a first dimension that accounts for as much variance as the second dimension several times (Prunas et al., 2012). It should be admitted that the one-factor model, which assumes that all items tap into a single dimension, showed a good fit in the two investigated samples of this study, too. A second-order factor might explain the co-variances of the primary factors without the need of discounting the informative value of the scores on the primary factors.

In some past studies a bifactor model of the SCL-90-R with nine orthogonally independent factors was successfully implemented in several adult samples (Urbán et al., 2014; Urbán et al., 2016; Arrindell et al., 2017), and in one adolescent sample (Rytilä-Manninen et al., 2016). This study confirmed the reproducibility of this bifactor structure of the SCL-90-R. It should be borne in mind that bifactor models can be very demanding in terms of sample size, particularly with a 90-item questionnaire and nine factors to fit. After successfully testing the bifactor model of the SCL-90-R in a Hungarian sample, Urbán et al. (2016) were unable to replicate the outcome when testing it in a Dutch sample that was supposed to be large enough (n = 1902) to allow the fitting. They reached only a marginal fit for the bifactor model in the Dutch sample when they allowed correlations between specific factors. However in a study carried out in a sample including 2593 patients with substance use and impulse control disorders, Arrindell et al. (2017) found support for both the standard bifactor structure of the SCL-90-R and the bifactor model with nine correlated specific factors. In this study, too, a bifactor implementation of the nine correlated specific factors showed a better fit than the standard bifactor model in both the urban and the rural samples. The bifactor model with oblique rotation of the nine specific factors showed a reasonable distribution of the explained common variance; the greater share was attributable to the general factor and the minimum load to the independent factors, with some exceptions.

As reported in a past study carried out in patients' samples (Urbán et al., 2016), the somatization, hostility, and phobic anxiety scales were linked to the dimensions that reflect their specific primary symptoms, especially in the urban sample. Overall, the hostility scale consistently showed relative independence from the general factor of distress in both samples, suggesting that it measures a dimension that fluctuates independently from the levels of psychological distress. This reminds of the empirically supported distinction between externalization and internalization problems in adolescents (Achenbach et al., 2017). However, a discussion on the topic and its related measurement issues (Achenbach et al., 2016) are beyond the scope of this study.

4.1. Research implications of the study

On a research ground, it is important to take into account that the mere observation of a good fit of a model is not a proof of its adequacy. Bifactor models tend to overfitting (Bonifay et al., 2016). Benchmarking indicators should be described to assure the model is acceptable. Whether or not each item fitted on its factor should be reported as well.

Another element worth noting is that the choice of a specific structural model depends on theoretical considerations, and not merely on the reaching of some thresholds in the parameters for fit estimation. In a bifactor model a general latent dimension contaminates the responses of the items, and it is very difficult to disentangle its effects from the influence of the group-factor specific dimensions. Therefore, the choice of a bifactor model of the SCL-90-R implies rejecting the use of its nine first-factors. The mere proof of the good fit of a model is not a demonstration of the existence of the hypothesized construct – in this case, a single general dimension of psychological distress underlying the response to the 90 items of the questionnaire. Indeed, structural models are specified and estimated in order to create a latent variable that investigates individual differences in psychological traits (i.e., in variables that cannot be directly observed). Structural models do not "discover" latent variables, they establish them (Bonifay et al., 2016, p. 2). In other words, proving that a door is unlocked - which is a prerequisite to enter a room - is conceptually distinct from proving that someone entered that room (Borsboom et al., 2004, p. 1062).

Recently a bias of fit indices towards the bifactor model over the hierarchical second-order model was highlighted in simulation studies (Morgan et al., 2015; Gignac, 2016). As a matter of fact, this study does not report a relevant increase in fit from the second-order implementation of the nine-factor model to its bifactor implementation. In the urban sample the change in the CFI, the TLI or the RMSEA was much smaller than the 0.01-0.02 conventionally accepted threshold for statistically relevant differences between models (Cheung and Rensvold, 2002; Chen, 2007). The fit of the models was also suspiciously good in the rural sample, suggesting that in this sample the power to distinguish between the hierarchical second-order and the bifactor model was inadequate despite the sample size being in the good-enough range for this kind of comparisons (Chen et al., 2006). There is some trade-off between the number of items in a scale and the number of observations, particularly for complex models such as the bifactor one. Therefore a sample size of 500 cases like the one of the rural sample might have been insufficient to implement a bifactor model of the SCL-90-R. Overall, the real reproducibility of a bifactor model of the SCL-90-R with nine independent factors remains an open issue for research.

4.2. Clinical implications of the study

The choice between the standard nine-factor structure of the SCL-90-R and its bifactor implementation has some consequences on the clinical use of the tool. The SCL-90-R is often used to test improvement over time in patients undergoing treatment (Kirchmann et al., 2011; Haagen et al., 2018). By focusing on the nine dimensions of psychological distress that are measured by the SCL-90-R, the therapist may identify those variables that are sensitive to change and those that are more resistant to it. Longitudinal studies are particularly helpful in establishing the predictive ability of the SCL-90-R, especially within the perspective of early intervention (Benarous et al., 2016; Raballo et al., 2017; Rice et al., 2018). There is some evidence that the scores on the scales of the SCL-90-R that tap into symptoms of depression and psychosis predict long-term persistence of subclinical psychotic symptoms (Rössler et al., 2007; Rössler et al., 2011a), and are associated with an enhanced risk of common mental disorders (Rössler et al., 2011b). Interventions aimed at the early detection of these constellations of symptoms and their treatment may reveal value to the purpose of preventing mental disorders and their psychological and social consequences.

The acceptable fit of the nine-factor structure of the SCL-90-R suggests it may be applied as a measure of global and specific distress in non-clinical samples. Particularly in adolescence, the SCL-90-R may be used to quantify the intensity of psychological distress and its distribution across nine common sets of symptoms, which may precede the onset of a major mental disorder by further aggregating in specific constellations (McGorry et al., 2018).

4.3. Strengths and limitations

The use of independent samples to test the reproducibility of the models and the application of adequate and robust methods of parameter estimation are the main strengths of the study. Besides the likely issue of statistical power in the rural sample, a great limitation of the study is the lack of a clinical sample for comparison. The samples that were investigated in this study might be different from clinical samples of adolescents. Therefore, the results of the measurement invariance CFA should not be generalized to samples with a different degree of psychological distress across the dimensions that are evaluated by the SCL-90-R. The factor structure of the SCL-90-R might not be necessarily the same for clinical and non-clinical samples. The measurement invariance across different levels of psychological distress of the factor structure of the SCL-90-R has been poorly investigated thus far, and it remains a topic of further testing and investigation.

5. Conclusion

This study provides some evidence in favor of the multidimensionality of the SCL-90-R, of the use of global severity indexes and of the specific domains in the measurement of psychological distress in community samples. Strong support for the bifactor implementation of the nine-factor structure of the SCL-90-R has been found, too. Hence the scores on the SCL-90-R would be used to assign the candidates to a single latent trait underlying their vulnerability to mental disorders, the persistence of the disorder after its onset, and the severity of symptoms (Caspi and Moffitt, 2018). Good reliability provides overall compelling evidence for its use in adolescent samples like those involved in this study. Because of its clarity and ease of administration, the SCL-90-R may be a particularly helpful measure of psychological distress in adolescents, especially in the perspective of longitudinal, repeated administration.

Compliance with ethical standards

[redacted]

Conflict of interest

No conflict of interest by any of the authors.

Role of funding source

Research funded by Università di Cagliari (2012 CAR - Contributo d'ateneo per la ricerca, on the share attributed to Dr Petretto). The funding body had no involvement in the design of the study, the collection, analysis and interpretation of data, the writing of the report and the decision to submit the article for publication.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.psychres.2018.12.146.

References

Achenbach, T.M., Ivanova, M.Y., Rescorla, L.A., 2017. Empirically based assessment and taxonomy of psychopathology for ages 1½-90 + years: developmental, multi-informant, and multicultural findings. Compr. Psychiatry 79, 4–18. https://doi.org/10.1016/j.comppsych.2017.03.006.

Achenbach, T.M., Ivanova, M.Y., Rescorla, L.A., Turner, L.V., Althoff, R.R., 2016. Internalizing/externalizing problems: review and recommendations for clinical and research applications. J. Am. Acad. Child Adolesc. Psychiatry 55 (8), 647–656. https://doi.org/10.1016/j.jaac.2016.05.012.

Adjaye-Gbewonyo, D., Rebok, G.W., Gallo, J.J., Gross, A.L., Underwood, C.R., 2018. Urbanicity of residence and depression among adults 50 years and older in Ghana and South Africa: an analysis of the WHO Study on Global AGEing and Adult Health (SAGE). Aging Ment. Health. 1–10. https://doi.org/10.1080/13607863.2018.1450839. 2018 Apr 10.

Arrindell, W.A., Barelds, D.P., Janssen, I.C., Buwalda, F.M., van der Ende, J., 2006. Invariance of SCL-90-R dimensions of symptom distress in patients with peri partum pelvic pain (PPPP) syndrome. Br. J. Clin. Psychol. 45, 377–391.

Arrindell, W.A., Urbán, R., Carrozzino, D., Bech, P., Demetrovics, Z., Roozen, H.G., 2017. SCL-90-R emotional distress ratings in substance use and impulse control disorders: one-factor, oblique first-order, higher-order, and bi-factor models compared. Psychiatry Res 255, 173–185.

Bekker, P.A., Merckens, A., Wansbeek, T.J., 1994. Identification, Equivalent Models, and Computer Algebra. Academic Press, San Diego.

Benarous, X., Consoli, A., Milhiet, V., Cohen, D., 2016. Early interventions for youths at high risk for bipolar disorder: a developmental approach. Eur. Child Adolesc. Psychiatry 25 (3), 217–233.

Bieda, A., Hirschfeld, G., Schönfeld, P., Brailovskaia, J., Zhang, X.C., Margraf, J., 2017. Universal happiness? Cross-cultural measurement invariance of scales assessing positive mental health. Psychol. Assess. 29 (4), 408–421. https://doi.org/10.1037/pas0000353.

Bonifay, W., Lane, S.P., Reise, SP., 2016. Three concerns with applying a bifactor model as a structure of psychopathology. Clin. Psychol. Sci. 5, 184–186.

Borsboom, D., Mellenbergh, G.J., van Heerden, J., 2004. The concept of validity. Psychol. Rev. 111 (4), 1061–1071.

Breslau, J., Marshall, G.N., Pincus, H.A., Brown, R.A., 2014. Are mental disorders more common in urban than rural areas of the United States? J. Psychiatr. Res. 56, 50–55. https://doi.org/10.1016/j.jpsychires.2014.05.004.

Browne, M.W., Cudeck, R., 1993. Alternative ways of assessing model fit. In: Bollen, KA, Long, JS (Eds.), Testing Structural Equation Models. Sage, Newbury Park, CA, pp. 136–161.

Byrne, B., van de Vijver, F.J.R., 2010. Testing for measurement and structural equivalence in large-scale cross-cultural studies: addressing the issue of nonequivalence. Int. J. Test. 10, 107–132.

Caspi, A., Moffitt, T.E., 2018. All for one and one for all: mental disorders in one dimension. Am. J. Psychiatry 175 (9), 831–844. https://doi.org/10.1176/appi.ajp.2018.17121383.

Chen, F.F., 2007. Sensitivity of goodness of fit indexes to lack of measurement invariance. Struct. Equ. Modeling 14, 464–504.

Chen, F.F., West, S.G., Sousa, K.H., 2006. A Comparison of bifactor and second-order models of quality of life. Multivariate Behav. Res. 41, 189–225.

Cheung, G.W., Rensvold, R.B., 2002. Evaluating goodness-of-fit indexes for testing measurement invariance. Struct. Equ. Modeling 9, 233–255.

Cyr, J.J., McKenna-Foley, J.M., Peacock, E., 1985. Factor structure of the SCL-90-R: is there one? J Pers. Assess. 49 (6), 571–578.

Dagani, J., Signorini, G., Nielssen, O., Bani, M., Pastore, A., Girolamo, G., et al., 2017. Meta-analysis of the Interval between the onset and management of bipolar disorder. Can. J. Psychiatry 62 (4), 247–258. https://doi.org/10.1177/0706743716656607.

Derogatis, L.R., Cleary, P.A., 1977. Confirmation of the dimensional structure of the SCL-90: a study in construct validation. J. Clin. Psychol. 33 (4), 981–989.

Derogatis, L.R., Savitz, K.L., 1999. The SCL-90-R, brief symptom inventory, and matching clinical rating scales. In: Maruish, M.E. (Ed.), The use of psychological testing for treatment planning and outcomes assessment. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US, pp. 679–724.

Derogatis, L.R., 1994. Symptom Checklist-90-R: administration, scoring and procedures manual, 3rd ed. National Computer Systems, Minneapolis.

Drancourt, N., Etain, B., Lajnef, M., Henry, C., Raust, A., Cochet, B., et al., 2013. Duration of untreated bipolar disorder: missed opportunities on the long road to optimal treatment. Acta Psychiatr. Scand. 127 (2), 136–144. https://doi.org/10.1111/j.1600-0447.2012.01917.x.

Dueber, D.M., 2016, December, December. Bifactor indices calculator: a Microsoft Excel-based tool to calculate various indices relevant to bifactor CFA models. Available at. http://sites.education.uky.edu/apslab/resources.

Eurostat., 2018. Your key to European Statistics. Methodology - Eurostat. https://ec.europa.eu/eurostat/web/rural-development/methodology Accessed on November, 2, 2018.

Flora, D.B., Curran, P.J., 2004. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. Psychol. Methods 9, 466–491.

Galobardes, B., Shaw, M., Lawlor, D.A., Lynch, J.W., Davey Smith, G., 2006. Indicators of socioeconomic position (part 1). J. Epidemiol. Community Health 60 (1), 7–12.

Geoffroy, P.A., Scott, J., 2017. Prodrome or risk syndrome: what's in a name? Int. J. Bipolar Disord. 5 (1), 7. https://doi.org/10.1186/s40345-017-0077-5.

Gignac, G.E., 2016. The higher-order model imposes a proportionality constraint: That is why the bifactor model tends to fit better. Intelligence 55, 57–68.

Gong, Y., Palmer, S., Gallacher, J., Marsden, T., Fone, D., 2016. A systematic review of the relationship between objective measurements of the urban environment and psychological distress. Environ. Int. 96, 48–57. https://doi.org/10.1016/j.envint.2016.08.019.

Haagen, J.F.G., van Rijn, A., Knipscheer, J.W., van der Aa, N., Kleber, R.J., 2018. The dissociative post-traumatic stress disorder (PTSD) subtype: a treatment outcome cohort study in veterans with PTSD. Br. J. Clin. Psychol. 57 (2), 203–222. https://doi.org/10.1111/bjc.12169.

Holzinger, K.J., Swineford, F., 1937. The bi-factor method. Psychometrika 2, 41–54.

Hu, L., Bentler, P.M., 1999. Cutoff criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. Struct. Equ. Modeling 6, 1–55.

Jöreskog, K.G., 1971. Simultaneous factor analysis in several populations. Psychometrika 36, 409–426.

Kessler, R.C., Amminger, G.P., Aguilar-Gaxiola, S., Alonso, J., Lee, S., Ustün, T.B., 2007. Age of onset of mental disorders: a review of recent literature. Curr. Opin. Psychiatry 20 (4), 359–364.

Kevin Chapman, L., Petrie, J., Vines, L., 2012. Factor structure analysis of the SCL-90-R in a community-based sample of African American women. Psychiatry Res. 199 (2), 140–144.

Kirchmann, H., Schreiber-Willnow, K., Seidler, K.P., Strauss, B., 2011. SCL-90-R symptom profile clusters among inpatients undergoing psychodynamic group psychotherapy: cluster stability, associations with clinical characteristics and treatment outcome. Psychopathology 44 (2), 71–82.

Mardia, K.V., 1970. Measures of multivariate skewness and kurtosis with applications. Biometrika 57, 519–530.

Marsh, H.W., Hau, K.T., Grayson, D., 2005. Goodness of fit in structural equation models. In: Maydeu-Olivares, A., McArdle, J.J. (Eds.), Contemporary psychometrics: A Festschrift for Roderick P. McDonald. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US, pp. 225–340.

Martin, G., Thomas, H., Andrews, T., Hasking, P., Scott, J.G., 2015. Psychotic experiences and psychological distress predict contemporaneous and future non-suicidal self-injury and suicide attempts in a sample of Australian school-based adolescents. Psychol. Med. 45 (2), 429–437. https://doi.org/10.1017/S0033291714001615.

McDonald, R.P., 1978. Generalizability in factorable domains: domain validity and generalizability: 1. Educ. Psychol. Meas. 38 (1), 75–79.

McGorry, P., van Os, J., 2013. Redeeming diagnosis in psychiatry: timing versus specificity. Lancet 381 (9863), 343–345. https://doi.org/10.1016/S0140-6736(12)61268-9.

McGorry, P.D., Hartmann, J.A., Spooner, R., Nelson, B., 2018. Beyond the "at risk mental state" concept: transitioning to transdiagnostic psychiatry. World Psychiatry 17 (2), 133–142. https://doi.org/10.1002/wps.20514.

Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., et al., 2010. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual. Life Res. 19, 539–549.

Morgan, G.B., Hodge, K.J., Wells, K.E., Watkins, M.W., 2015. Are Fit indices biased in favor of bi-factor models in cognitive ability research?: a comparison of fit in correlated factors, higher-order, and bi-factor models via Monte Carlo simulations. J. Intell. 3 (1), 2–20. https://doi.org/10.3390/jintelligence3010002.

Narayanan, A., 2012. A review of eight software packages for structural equation modelling. Am. Stat. 66, 129–138.

Nelson, B., McGorry, P.D., Wichers, M., Wigman, J.T.W., Hartmann, J.A., 2017. Moving from static to dynamic models of the onset of mental disorder: a review. JAMA Psychiatry 74 (5), 528–534. https://doi.org/10.1001/jamapsychiatry.2017.0001.

Oliver, D., Davies, C., Crossland, G., Lim, S., Gifford, G., McGuire, P., et al., 2018. Can we reduce the duration of untreated psychosis? A systematic review and meta-analysis of controlled interventional studies. Schizophr. Bull. 44 (6), 1362–1372. https://doi.org/10.1093/schbul/sbx166.

Peen, J., Schoevers, R.A., Beekman, A.T., Dekker, J., 2010. The current status of urban-rural differences in psychiatric disorders. Acta Psychiatr. Scand. 121 (2), 84–93. https://doi.org/10.1111/j.1600-0447.2009.01438.x.

Penttilä, M., Jääskeläinen, E., Hirvonen, N., Isohanni, M., Miettunen, J., 2014. Duration of untreated psychosis as predictor of long-term outcome in schizophrenia: systematic review and meta-analysis. Br. J. Psychiatry 205 (2), 88–94.

Pornprasertmanit, S., Miller, P., Schoemann, A., Rosseel, Y., 2013. semTools: useful tools for structural equation modeling. R package version 0.3-2. http://CRAN.R-project.org/package=semTools.

Preti, A., Siddi, S., Vellante, M., Scanu, R., Muratore, T., Gabrielli, M., et al., 2015. Bifactor structure of the schizotypal personality questionnaire (SPQ). Psychiatry Res. 230 (3), 940–950.

Prunas, A., Sarno, I., Preti, E., Madeddu, F., Perugini, M., 2012. Psychometric properties of the Italian version of the SCL-90-R: a study on a large community sample. Eur. Psychiatry 27 (8), 591–597.

Core Team., R, 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria URL. http://www.Rproject.org/.

Raballo, A., Poletti, M., McGorry, P., 2017. Architecture of change: rethinking child and adolescent mental health. Lancet Psychiatry 4 (9), 656–658.

Reise, S.P., 2012. The rediscovery of bifactor measurement models. Multivariate Behav. Res. 47, 667–696.

Reise, S.P., Bonifay, W.E., Haviland, M.G., 2013. Scoring and modeling psychological measures in the presence of multidimensionality. J. Pers. Assess. 95 (2), 129–140.

Reise, S.P., Moore, T.M., Haviland, M.G., 2010. Bifactor models and rotations: exploring the extent to which multidimensional data yield univocal scale scores. J. Person. Assess. 92, 544–559.

Rice, S.M., Purcell, R., McGorry, PD., 2018. Adolescent and young adult male mental health: transforming system failures into proactive models of engagement. J. Adolesc. Health 62 (3S), S9–S17.

Rodriguez, A., Reise, S.P., Haviland, M.G., 2016. Applying bifactor statistical indices in the evaluation of psychological measures. J. Person. Assess. 98 (3), 223–237.

Rosseel, Y., 2012. lavaan: an R package for structural equation modeling. J. Stat. Softw. 48, 1–36. URL. http://www.jstatsoft.org/v48/i02/.

Rössler, W., Angst, J., Gamma, A., Haker, H., Stulz, N., Merikangas, K.R., et al., 2011a. Reappraisal of the interplay between psychosis and depression symptoms in the pathogenesis of psychotic syndromes: results from a twenty-year prospective community study. Eur. Arch. Psychiatry Clin. Neurosci. 261 (1), 11–19.

Rössler, W., Hengartner, M.P., Ajdacic-Gross, V., Haker, H., Gamma, A., Angst, J., 2011b. Sub-clinical psychosis symptoms in young adults are risk factors for subsequent common mental disorders. Schizophr. Res. 131 (1-3), 18–23.

Rössler, W., Riecher-Rössler, A., Angst, J., Murray, R., Gamma, A., Eich, D., et al., 2007. Psychotic experiences in the general population: a twenty-year prospective community study. Schizophr. Res. 92 (1-3), 1–14.

Rytilä-Manninen, M., Fröjd, S., Haravuori, H., Lindberg, N., Marttunen, M., Kettunen, K., et al., 2016. . Psychometric properties of the symptom checklist-90 in adolescent psychiatric inpatients and age- and gender-matched community youth. Child Adolesc. Psychiatry Ment. Health 10, 23.

Santesteban-Echarri, O., Paino, M., Rice, S., González-Blanch, C., McGorry, P., Gleeson, J., et al., 2017. . Predictors of functional recovery in first-episode psychosis: a systematic review and meta-analysis of longitudinal studies. Clin. Psychol. Rev. 58, 59–75. https://doi.org/10.1016/j.cpr.2017.09.007.

Savill, M., D'Ambrosio, J., Cannon, T.D., Loewy, R.L., 2018. Psychosis risk screening in different populations using the prodromal questionnaire: a systematic review. Early Interv. Psychiatry 12 (1), 3–14. https://doi.org/10.1111/eip.12446.

Schmid, J., Leiman, L.J., 1957. The development of hierarchical factor solutions. Psychometrika 22, 53–61.

Schultze-Lutter, F., Michel, C., Schmidt, S.J., Schimmelmann, B.G., Maric, N.P., Salokangas, R.K., et al., 2015. . EPA guidance on the early detection of clinical high risk states of psychoses. Eur. Psychiatry 30 (3), 405–416. https://doi.org/10.1016/j.eurpsy.2015.01.010.

Sereda, Y., Dembitskyi, S., 2016. Validity assessment of the symptom checklist SCL-90-R and shortened versions for the general population in Ukraine. BMC Psychiatry 16, 300. https://doi.org/10.1186/s12888-016-1014-3.

Sijtsma, K., 2009. On the use, the misuse, and the very limited usefulness of Cronbach's alpha. Psychometrika 74, 107–120.

Stickley, A., Koyanagi, A., Roberts, B., McKee, M., 2015. Urban-rural differences in psychological distress in nine countries of the former Soviet Union. J. Affect. Disord. 178, 142–148. https://doi.org/10.1016/j.jad.2015.02.020.

Stochl, J., Khandaker, G.M., Lewis, G., Perez, J., Goodyer, I.M., Zammit, S., et al., 2015. Mood, anxiety and psychotic phenomena measure a common psychopathological factor. Psychol. Med. 45 (7), 1483–1493. https://doi.org/10.1017/S003329171400261X.

Tomioka, M., Shimura, M., Hidaka, M., Kubo, C., 2008. The reliability and validity of a Japanese version of symptom checklist 90 revised. Biopsychosoc. Med. 2, 19. https://doi.org/10.1186/1751-0759-2-19.

Urbán, R., Arrindell, W.A., Demetrovics, Z., Unoka, Z., Timman, R., 2016. Cross-cultural confirmation of bi-factor models of a symptom distress measure: symptom checklist-90-revised in clinical samples. Psychiatry Res. 239, 265–274.

Urbán, R., Kun, B., Farkas, J., Paksi, B., Kökönyei, G., Unoka, Z., et al., 2014. Bifactor structural model of symptom checklists: SCL-90-R and brief symptom inventory (BSI) in a non-clinical community sample. Psychiatry Res. 216 (1), 146–154.

Vandenberg, R.J., Lance, C.E., 2000. A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. Organ. Res. Methods 3, 4–70.

Vassend, O., Skrondal, A., 1999. The problem of structural indeterminacy in multidimensional symptom report instruments. The case of SCL-90-R. Behav. Res. Ther. 37 (7), 685–701.

Vassos, E., Agerbo, E., Mors, O., Pedersen, C.B., 2016. Urban-rural differences in incidence rates of psychiatric disorders in Denmark. Br. J. Psychiatry. 208 (5), 435–440. https://doi.org/10.1192/bjp.bp.114.161091.

Veenhoven, R., 1996. Developments in satisfaction research. Soc. Indic. Res. 37, 1–46. https://doi.org/10.1007/BF00300268.

World Medical Association, 2013. Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 310, 2191–2194.