

Factor Structure of the Sussex-Oxford Compassion Scales

Júlia Halamová¹ and Martin Kanovský²

¹Institute of Applied Psychology, Faculty of Social and Economic Sciences, Comenius University in Bratislava, Bratislava, Slovakia

²Institute of Social Anthropology, Faculty of Social and Economic Sciences, Comenius University in Bratislava, Bratislava, Slovakia

Abstract

Without a reliable and valid instrument of compassion and self-compassion, it is hard to conduct quality research in compassion even though it is a rapidly growing area of interest in many scientific fields. However, recently the Sussex-Oxford Compassion Scales (SOCS; Gu et al., 2020) were developed, with two parallel versions for compassion (SCOS-O) and self-compassion (SCOS-S) containing 20 items each. Because the SOCS scales are newly developed measures of compassion, this is the first study to be conducted beyond the original research (Gu et al., 2020) that validates their factor structure. Our sample consisted of 1080 respondents; 19.35% were males and 80.65% were females. Mean age was 29.29 ($SD = 11.15$). Convenience sampling via social networks was used. To analyse the data, we reproduced all the confirmatory models in Gu et al. (2020) for both scales: the one-factor model, five-factor model, and five-factor hierarchical model. Our research results showed that for the SCOS-O, the multidimensional definition of compassion for others fits the factor structure of the scale, but this does not apply to the SCOS-S. Self-compassion has two dominant factors over and above five specific factors: Rational Compassion (containing two specific factors: Recognising Suffering and Understanding the Universality of Suffering) and Emotional/Behavioural Compassion (containing three specific factors: Feeling for the Person Suffering, Tolerating Uncomfortable Feelings, and Acting or Being Motivated to Act to Alleviate Suffering). Therefore, the total self-compassion score is unsuitable for use because the scale lacks essential unidimensionality; however, the compassion for others total score can be used safely.

Keywords: compassion, self-Compassion, psychometric analysis, factor analysis, assessment

✉ Julia Halamová, Institute of Applied Psychology, Faculty of Social and Economic Sciences, Comenius University in Bratislava. Mlynské luhy 4, 821 05 Bratislava, Slovakia. E-mail: julia.halamova@gmail.com

Acknowledgement

This work was supported by the Slovak Research and Development Agency under the Contract no. PP-COVID-20-0074. Writing this work was supported by the Vedecká grantová agentúra VEGA under Grant 1/0075/19.

We would like to acknowledge Martina Baránková, Bronislava Strnádelová, and Simona Tittelová for help with back translation and Mária Poláková, Veronika Špaňová, and Simona Tittelová for the help with data gathering.

Introduction

Compassion is becoming a growing area of interest in psychology, medicine, healthcare, organisational science, education, justice, and social work (Seppälä et al., 2017) so it is important we have a reliable and valid instrument to measure compassion. The lack of a reliable and valid instrument affects the quality of research in this rapidly expanding area. As Strauss et al. (2016, p. 25) conclude in their review of the existing scales of compassion: “no scale exists that comprehensively measures compassion and provides scores with acceptable levels of reliability and validity”. This is also true of the newest scale of compassion, the Compassionate Engagement and Action Scales (CEAS; Gilbert et al., 2017), which was not included in the review by Strauss et al. (2016). According to Halamová et al. (2021), the Compassion to Others and Compassion from Others scales of the CEAS have a good factor structure and psychometric properties, but the total score of the Compassion for Self scale of the CEAS cannot be reliably used.

The difficulty of performing research in this area may also be related to the ongoing dispute over the definition of compassion (Strauss et al., 2016). Compassion is defined differently by different researchers: Ekman (2014) perceives compassion as the experience of empathic distress, Gilbert et al. (2017) define it as an awareness of suffering associated with motivation and helping behaviours aimed at reducing it, Sprecher and Fehr (2005) see compassion as a special kind of love, and Goetz et al. (2010) consider compassion a distinct emotional state. By contrast, Jazaieri et al. (2013) and Strauss et al. (2016) define compassion as a multidimensional construct. Strauss et al. (2016, p. 19) have formulated a definition of compassion on the basis of a systematic review and by consolidating various previous definitions: “1) Recognizing suffering; 2) Understanding the universality of suffering in human experience; 3) Feeling empathy for the person suffering and connecting with the distress (emotional resonance); 4) Tolerating uncomfortable feelings aroused in response to the suffering person (e.g. distress, anger, fear) so remaining open to and accepting of the person suffering; and 5) Motivation to act/acting to alleviate suffering.” The five-item definition by Strauss et al. (2016) was further tested and supported by Gu et al. (2016), who conducted a series of factor analyses of items from the existing self-report compassion measures to detect the underlying structure of compassion.

Development of Sussex-Oxford Compassion Scales

Based on this theoretically and empirically supported definition of compassion, a new instrument has been developed: the Sussex-Oxford Compassion Scales (SOCS; Gu et al., 2020). It is so new it has been used in only one research study (Benda & Vyhnánek, 2019). The scale was created with two corresponding versions for compassion and self-compassion – the Sussex-Oxford Compassion for Others Scale (SOCS-O) and the Sussex-Oxford Compassion for the Self Scale (SOCS-S) –

the authors assume that both constructs have the same basis and so do not differentiate between them (e.g., Gilbert, 2014; Gu et al., 2020; Strauss et al., 2016).

There were four stages in the development of SOCS-O and SOCS-S (Gu et al., 2020). In the first stage, the authors generated and reviewed the items with the support of experts and non-experts (Gu et al., 2020). In the second stage, the number of items was reduced. In the third, the psychometric properties were tested, and the factor structure of the scales was assessed using a confirmatory factor analysis. In the fourth stage, the factor structures of both versions were cross-validated (Gu et al., 2020).

Reliability of Sussex-Oxford Compassion Scales

Gu et al. (2020) tested the internal consistency of SOCS using Cronbach's alpha coefficients and total omega. As Gu et al. (2020) report, coefficients ranging between .61 and .97 are at the outer limit of the recommended values. The original study did not measure test-retest reliability (Gu et al., 2020). According to Tavakol and Dennick (2011), for good internal consistency the values should be more than .70 and less than .90. If the consistency coefficient is higher than .95 there is a risk of redundant items and the scale being too long (Tavakol & Dennick, 2011). On the other hand, some authors consider values under .70 to be acceptable in psychological research (Kline, 1999).

Validity and the Factor Structure of Sussex-Oxford Compassion Scales

The convergent and discriminant validity of SOCS-O and SOCS-S was analysed, and more than three quarters of the correlations were as expected (Gu et al., 2020). Generally, SOCS-O had large significant positive correlations with the Santa Clara Brief Compassion Scale (SCBCS; Hwang et al., 2008), while SOCS-S had large significant positive correlations with the Self-Compassion Scale – Short Form (SCS-12; Raes et al., 2011). Related to discriminant validity and previous research on negative correlation of self-compassion and psychopathology (e.g. MacBeth, & Gumley, 2012), significant negative correlations were found between SOCS-S and psychopathology symptoms. There were negative correlations between SOCS-S and depersonalisation with -.36 for SOCS-S and -.24 for SOCS-O, and the emotional exhaustion subscales with -.34 for SOCS-S and -.10 for the SOCS-O of the Maslach Burnout Inventory–Human Services Survey (MBIHSS; Maslach et al., 1981) and between SOCS-O and the personal distress subscale of the Interpersonal Reactivity Index (IRI; Davis, 1980) with -.24 for SOCS-S and -.16 for SOCS-O. In addition, both forms of SOCS correlated significantly and positively with the 15-item version of the Five-Facet Mindfulness Questionnaire (FFMQ-15; Baer et al., 2012) with .55 for SOCS-S and .26 for SOCS-O, the Short Warwick-Edinburgh Mental Well-Being Scale (SWEMWBS; Stewart-Brown et al., 2009) with .57 for SOCS-S and .24 for SOCS-O, and negatively with the Depression, Anxiety, and

Stress Scale – Short Form (DASS; Henry & Crawford, 2005) varying between $-.36$ to $-.45$ for SOCS-S and between $-.03$ to $-.07$ for SOCS-O. Unfortunately, the original study did not include an instrument for measuring self-criticism (Gu et al., 2020), even though self-compassion is considered to be the antidote to self-criticism (Neff, 2003b) as self-compassionate interventions lower the level of self-criticism. It would therefore have been useful to be able to test the discriminant validity of the scales, not just in relation to self-criticism, but in general as well.

For SOCS, a confirmatory factor analysis was conducted, first on a sample of health care staff and afterwards on university students for the purposes of cross-validation (Gu et al., 2020). The SOCS-O factor analysis and cross-validation of factor analysis both showed best fit for the five-factor hierarchical model related to the five-item definition. For SOCS-S data, the five-factor hierarchical model had an adequate fit in the validation as well as the cross-validation samples.

Aim of the Study

No other study, besides the original study, has been published that analyses the factor structure of the newly developed Sussex-Oxford Compassion Scales (Gu et al., 2020). In addition, this is the first study to examine the SOCS factor structure in an international sample. Therefore, our goal was to translate SOCS into Slovak and analyse the factor structures of the two scales, SOCS-O and SOCS-S. Our hypothesis was that there is a single general overarching factor above the five subfactors in the Slovak sample, which is similar to the findings of Gu et al. (2020) regarding the English sample.

Methods

Participants and Procedure

Our sample consisted of 1080 respondents; out of these 209 (19.35%) were males and 871 (80.65%) were females. Mean age was 29.29 ($SD = 11.15$) ranging from 18 to 75 years. Out of these, 409 reported working in helping professions, 157 in non-helping professions and 514 did not report a profession. We gathered the data online via various social networks. Convenience sampling was used. To calculate the required number of participants, we used the A-Priori Sample Size Calculator for Structural Equation Models (Soper, 2020), that would give us the desired statistical power of $.80$, probability level of $.05$, number of latent variables 6, number of observed variables 20, anticipated effect size: $.15$. All the participants signed an online informed consent form prior to participating in the research study. Data were collected in accordance with the ethical standards of the related university ethical committee and in accordance with the 1964 Helsinki declaration and its later amendments.

Measures

The Sussex-Oxford Compassion Scales (SOCS; Gu et al. 2020) was developed with two corresponding versions: for compassion – the Sussex-Oxford Compassion for Others Scale (SOCS-O); and for self-compassion – the Sussex-Oxford Compassion for the Self Scale (SOCS-S). Both scales consist of five subscales: Recognising Suffering, Understanding the Universality of Suffering, Feeling for the Person Suffering, Tolerating Uncomfortable Feelings, and Acting or Being Motivated to Act to Alleviate Suffering. Each subscale consists of four items. The scales use a five-point Likert response format (1 = *not at all true*, 2 = *rarely true*, 3 = *sometimes true*, 4 = *often true*, 5 = *always true*). For SOCS-O, the Recognising Suffering subscale contains for example the following item, “I recognise when other people are feeling distressed without them having to tell me.” and for SOCS-S, “I’m good at recognising when I’m feeling distressed.” For SOCS-O, the Tolerating Uncomfortable Feelings subscale includes for example the item, “I stay with and listen to other people when they’re upset even if it’s hard to bear.” and for SOCS-S, “I connect with my own distress without letting it overwhelm me.” For SOCS-O, the Feeling for the Person Suffering subscale comprises items such as “When someone is going through a difficult time, I feel kindly towards them.” and for SOCS-S, “When I’m going through a difficult time, I feel kindly towards myself.” Both SOCS-O (“I know that we can all feel upset at times when we are wronged.”) and SOCS-S (“I know that we can all feel distressed when things don’t go well in our lives.”) contain similarly formulated items for Understanding the Universality of Suffering. For SOCS-O, Acting or Being Motivated to Act to Alleviate Suffering contains for example the item “When I see someone in need, I try to do what’s best for them.” and for SOCS-S “When I’m upset, I try to do what’s best for myself.” Both the SOCS-O and SOCS-S scales were back-translated from English to Slovak and then back from Slovak to English and compared by a research team consisting of psychologists and in consultation with a translator.

Data Analysis

All statistical analyses were conducted in Mplus version 7.4 (Muthén & Muthén, 1998-2010). Our analytical strategy was to reproduce all the confirmatory models from Gu et al. (2020) for both scales (SOCS-O and SOCS-S): one-factor model, five-factor model, five-factor hierarchical model (in which all five factors load on a general compassion factor). We reported the same indices of fit as Gu et al. (2020): χ^2 and degrees of freedom (*df*), Comparative Fit Index (CFI), Non-normed Fit Index (NNFI/TLI), Standardised Root Mean Square Residuals (SRMR), Root Mean Square Error of Approximation (RMSEA) with 90% confidence intervals, and Akaike Information Criterion (AIC). Following Williams et al. (2014) and Gu et al. (2020), we used both liberal and conservative cut-off points for acceptable fit for the CFI (greater than .90 as liberal; greater than .95 as conservative), RMSEA (less than .10 as liberal; less than .05 as conservative), NNFI (greater than .90 as liberal; greater

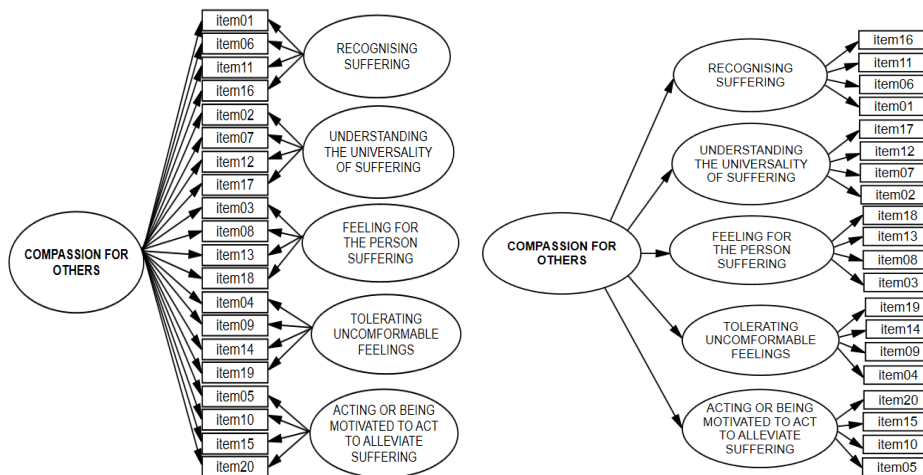
than .95 as conservative), and SRMR (less than .10 as liberal; less than .05 as conservative). The same estimator (maximum likelihood with robust correction, MLR) was used. We compared the results and discussed some of the implications. We fitted two more confirmatory models, namely bifactor models (see Figure 1 below). We agree that the higher-order factor models used by Gu et al. (2020) represent a common approach to capturing construct-relevant multidimensionality (but see Cucina & Byle, 2016). However, this approach relies on a restrictive implicit assumption – a proportionality constraint: the ratio of the variance attributed to the higher-order general factor versus uniquely attributed to the first-order factor is a constant for all items associated with a single first-order factor (Reise, 2012). Even more importantly, the higher-order general factor does not explain any additional variance besides the variance already explained by the first-order factors. On the other hand, bifactor models (Reise, 2012; see Figure 1 below) allow for the estimation of an overarching general construct without relying on the restrictive implicit assumption of the proportionality constraint, and for a separate and straightforward assessment of variance uniquely attributable to specific and global overarching factors. In this bifactor approach, the covariance among a set of items can be explained by a set of orthogonal factors including one Global (G) general overarching factor and several orthogonal Specific (S) factors. As each item is used to simultaneously load on the G-factor and one S-factor, the covariance is divided into a G-factor underlying all the items, and S-factors corresponding to the covariance not explained by the G-factor. As such, the G-factor estimated as part of such a bifactor model provides us with a direct way of testing for the presence of a global general overarching construct underlying responses to all items, while also acknowledging that important distinctions exist at the subscale level (expressed by specific S-factors).

A couple of statistical indices for a single common general factor have been proposed by Rodriguez et al. (2016): hierarchical omega ω_h (the degree to which composite scale scores are interpretable as a measure of a single common general factor, and specific factors and errors are treated as a nuisance), and ECV (explained common variance: the ratio of variance explained by the general factor divided by the variance explained by the general plus the specific factors, therefore only the relation between a general factor and specific factors is captured). Note that ECV is a measure of essential unidimensionality (high ECV values mean the construct is essentially unidimensional in spite of some multidimensionality – specific factors contribute a minimal amount of explained variance), but it tells us nothing about the strength of the general factor: a single weak factor could be unidimensional, but the raw scores could reflect a lot of errors. By contrast, high hierarchical omega values mean that something general and strong captures a lot of the variance, but we do not know if this strong source is a single source or not. Therefore, when relying on raw scores, it is necessary to obtain a scale where both measures have high values (at least .70 for hierarchical omega, and at least .85 for ECV, see Stucky & Edelen [2014]). It should be emphasised that the statistical significance of the factor loadings

on the general factor in the hierarchical model does not test the essential unidimensionality of a construct (which is required if the unit-weighted total score is recommended).

Figure 1

Bifactor Model (Left Panel) and Five-Factor Hierarchical Model (Right Panel) for The Compassion for Others Scale



Note. Residual variances of items and disturbances of group factors are not displayed.

We reported all relevant reliability indices for both scales (Compassion for Others, Compassion for the Self) and their subscales, including Cronbach’s alpha and McDonald’s omega total (these indices are reported in Gu et al., 2020), hierarchical omega and Explained Common Variance (ECV; which were not reported by Gu et al., 2020). Furthermore, we inspected another valuable source of statistical information provided in Gu et al. (2020) and derived from our data: factor loadings of general factors in hierarchical models, factor loadings of general factors in bifactor models, and magnitude of correlations among subscales. Gu et al. (2020) did not perform chi-square difference or likelihood-ratio model comparisons, therefore we also refrain from doing so.

Results

The fit indices of all the models tested are shown in Table 1. We used the same statistical models as Gu et al. (2020, p. 11) and added bifactor models.

Table 1*Fit Indices for all Fitted Compassion Models*

Scale	Model	CFI	RMSEA [90% CI]	NNFI	SRMR	Satorra- Bentler χ^2 (df)	AIC
Compassion for Others	One-factor	.758	.095 [.092, .098]	.729	.083	1833.089 (170)	47960.957
	Five-factor	.949	.046 [.042, .049]	.939	.043	520.122 (160)	45748.768
	Five-factor hierarchical ^a	.943	.047 [.044, .051]	.935	.050	565.618 (165)	45811.104
	Bifactor	.962	.042 [.042-.044]	.944	.047	434.709 (150)	
Compassion for the Self	One-factor	.632	.112 [.109, .115]	.589	.120	2470.422 (170)	56363.297
	Five-factor	.923	.053 [.048, .057]	.909	.052	638.761 (160)	53916.914
	Five-factor hierarchical ^a	.915	.059 [.055, .063]	.900	.078	786.631 (165)	54101.750
	Bifactor	.945	.051 [.049, .053]	.931	.050	569.801 (150)	53953.042

Note. AIC = Akaike Information Criterion; CFI = Comparative Fit Index; CI = Confidence Interval; NNFI = Non-normed fit index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardised Root Mean Square Residual. Bold indices indicate acceptable fit according to liberal cut-off criteria. ^aFive-factor hierarchical stands for a model in which all five factors load on an overarching compassion factor.

We compared our results with those in Gu et al. (2020), considering only the health care staff samples: these samples are larger – 1242 and 1216 – and therefore more similar to our sample size – 1080. Some indices of fit are very sensitive to sample size so it would not be appropriate to compare our sample with the student samples (371).

First, we can conclude that our results almost exactly replicate the results of Gu et al. (2020), as can be seen in Table 1 (the values are of course slightly different, but the pattern is the same), and Table 1 in Gu et al. (2020, p. 11). However, we have some reservations about following their conclusion: “Based on both the fit indices and significance of factor loadings, the five-factor hierarchical model can be interpreted as best fitting the data.” (Gu et al., 2020, p. 11). The fit of the five-factor hierarchical model is consistently worse than the fit of the five-factor model in the RMSEA and SRMR fit indices, and it is worse in the AIC criteria, both in Gu et al. (2020) and in our analyses. We admit that the fit is slightly better or the same in the CFI and NNFI indices, but this improvement is limited to the third decimal place. Moreover, Gu et al. (2020) did not perform the model comparison (chi-square difference tests, or likelihood-ratio test), and they based their interpretation solely on comparing fit

indices and on the significance of the factor loadings: “Based on both the fit indices and significance of factor loadings, the five-factor hierarchical model can be interpreted as best fitting the data” (p. 10). We cannot calculate the Satorra-Bentler scaled chi-square difference test, nor the Satorra-Bentler scaled likelihood-ratio test for their models, because they do not report the values of the Satorra-Bentler scaling correction factor (see Satorra, [2000], Satorra & Bentler, [2010]). Our Table 1 strictly reproduces their Table 1, except for the addition of the bifactor models. We also did a statistical model comparison (Table 2). In terms of the statistical chi-square testing of nested models, the five-factor hierarchical model had a significantly better fit than the one-factor model in both scales. However, its fit is significantly worse than the five-factor correlated model in both scales. On the other hand, the bifactor model significantly outperforms the five-factor correlated model in both scales.

Table 2

Statistical Comparison of the Models

Scale	Model	χ^2 (df)	<i>p</i>
Compassion for Others	One-factor	-	-
	Five-factor hierarchical ^a	654.023 (5)	< .001
	Five-factor	47.849 (5)	< .001
	Bifactor	83.130 (10)	< .001
Compassion for Self	One-factor	-	-
	Five-factor hierarchical ^a	3011.006 (5)	< .001
	Five-factor	133.055 (5)	< .001
	Bifactor	69.069 (10)	< .001

Note. ^aFive-factor hierarchical stands for a model in which all five factors load on an overarching compassion factor.

We fully agree with Gu et al. (2020) that indices of fit should not be overestimated when deciding which particular model to select and that factor intercorrelations and loadings should be considered as well. But we disagree that the significance alone should be closely and carefully inspected: a close and careful inspection should be done for the magnitude loadings as well. With a large sample size, a loading of .20 or even less could well be significant (e.g. Stevens, 2009, p. 332). It seems that Gu et al. (2020) are inclined to think that if all factor loadings of the general factor in the hierarchical model are statistically significant, then that model should be preferred despite the factor loadings having a very small magnitude. One should carefully inspect both the statistical significance and the magnitudes of the factor loadings to detect the possible source of the misfit. For example, if the sample size is large, the factor loadings could all be small, but still significant. Even if the loadings were large enough (and of course significant as a consequence), their magnitudes could display a very unbalanced pattern indicating that a single general overarching factor did not capture the variance in specific factors very well. In other words, it could cast serious doubt on the essential unidimensionality of the scale. As

we will see presently, this occurred both in Gu et al. (2020) and in our data for the Compassion for the Self scale. Table 3 reports (standardised) factor loadings to an overall compassion factor in the five-factor hierarchical models in Gu et al. (2020, Table S4, Electronic Supplementary Materials) and our data.

Table 3

Standardised Loadings of Factors to an Overall Compassion Factor in The Five-Factor Hierarchical Models

	Compassion for Others		Compassion for the Self	
	Gu et al. (2020)	Present article	Gu et al. (2020)	Present article
Recognising suffering	.78 (0.02)*	.66 (0.03)*	.59 (0.03)*	.38 (0.04)*
Understanding the universality of suffering	.58 (0.03)*	.65 (0.03)*	.34 (0.03)*	.43 (0.03)*
Feeling for the person suffering	.98 (0.01)*	.99 (0.01)*	.99 (0.01)*	.95 (0.02)*
Tolerating uncomfortable feelings	.93 (0.02)*	.99 (0.01)*	.94 (0.01)*	.90 (0.03)*
Acting or being motivated to act to alleviate suffering	.92 (0.01)*	.97 (0.01)*	.94 (0.01)*	.94 (0.02)*

Note. Standard errors are given in parentheses. * $p < .001$.

It is immediately clear from the results presented in Table 3 that (1) the pattern of factor loadings in Gu et al. (2020) has been replicated; (2) all the factor loadings are significant and have high values; (3) the pattern of the magnitudes of factor loadings is highly asymmetrical in both samples: the factor loadings of the first two specific factors (Recognising Suffering and Understanding the Universality of Suffering) on the general factor are substantially lower in both datasets and both scales (the Compassion for Others scale being less affected in both datasets).

There are other indications that this hypothesis that there are two general dominant factors instead of a single general overarching factor could be true (or at least deserves serious testing). Gu et al. (2020, p. 11) note that “all factor intercorrelations in the five-factor model were significant”, but unfortunately they did not report them. A careful inspection of the factor loadings of all the models (including bifactor models and five-factor hierarchical models) usually reveals very useful information about the underlying structure of the construct (see Figure 1). Asymmetry in the factor loadings of the general factor could be detected in the factor intercorrelations as well: if there were two dominant factors instead of one, we would expect the intercorrelations of specific factors within these two dominant domains to be of a greater magnitude than the intercorrelations of specific factors across these two dominant domains. This same pattern should be detected among the intercorrelations of the total subscale scores (they are reported in Gu et al., 2020, Table S4, Supplementary Materials).

For the Compassion for Others scale, the correlation coefficients within the two dominant domains are on average .68 and the correlation coefficients across the two general domains are on average .53 in our data – that is not a large difference. On the other hand, for the Compassion for the Self scale, the correlation coefficients within the two dominant domains are on average .63 and the correlation coefficients across two dominant domains are on average only .29 in our data – that is a very large difference in magnitude. When we inspect the intercorrelations of the latent factors in the five-factor models in our data (Supplementary Materials, Tables S1 and S2), this pattern is even stronger: .87/.64 for the Compassion of Others scale, and .80/.35 for the Compassion of the Self scale. This comes as no surprise because the five-factor correlated models are nested in the five-factor hierarchical models. The same pattern is clear in the Gu et al. (2020) data: .64/.56 for the Compassion of Others scale, and .68/.40 for the Compassion of the Self scale. All the values are presented in detail in the Supplementary Materials, Tables S1 and S2. We can conclude that the evidence that there is a single, overarching general factor for both scales is not convincingly strong, especially for the Compassion for the Self scale.

It is worth repeating here that the statistical significance of the factor loadings on the general factor in the hierarchical model has little to do with the essential unidimensionality of the construct. The coefficients alpha and omega total, reported by Gu et al. (2020), are not tests of the essential unidimensionality either. It is unclear why Gu et al. (2020) used the one-factor omega total calculated using McNeish's (2018) Excel spreadsheet (this omega total index just relaxes the assumption of tau-equivalence) and not the omega hierarchical recommended by McNeish (2018, p. 417) in cases where the "researchers believe that the items in the scale are organized in hierarchical factors". Omega hierarchical (and the ECV index mentioned earlier) directly assess the strength of the general factor and its essential unidimensionality, but they require theoretical justification and careful modelling that specifies the structure of group factors (see e.g. Raykov & Marcoulides, 2019; Savalei & Reise, 2019 for critical reviews of McNeish's approach and useful recommendations). In the present case, Gu et al. (2020) provide a theoretical explanation for assuming a five-factor structure and a strong general factor, so we fit two confirmatory bifactor models (see Figure 1) with five specific group factors. We subsequently calculate the omega total, omega hierarchical and ECV indices based on the factor loadings (and residual variances) of these bifactor models, as recommended in the psychometric literature (e.g. Rodriguez et al., 2016).

Not surprisingly, the fit indices of the bifactor models were better than the five-factor hierarchical models for both Compassion for Others and Compassion for the Self (see Table 1 above). We can therefore calculate the reliability indices and inspect the factor loadings. All indices are presented in Table 4.

Table 4*Reliability and ECV Coefficients for SOCS-O and SOCS-S Scale and Subscale Items*

	Compassion for others				Compassion for the self			
	Alpha	Omega total	Omega hierarchical	ECV	Alpha	Omega total	Omega hierarchical	ECV
Total scale	.93	.96	.90	.78	.89	.91	.78	.53
1	.87	-	-	-	.80	-	-	-
2	.88	-	-	-	.82	-	-	-
3	.75	-	-	-	.72	-	-	-
4	.74	-	-	-	.69	-	-	-
5	.85	-	-	-	.80	-	-	-

Note. ECV = Explained Common Variance.

We can see that the Omega total values are very high for both scales (these results replicate the results in Gu et al. [2020]). For the Compassion for Others scale, the Omega hierarchical is high: 90% of total variance is explained by a single general factor. However, the ECV value is .78, which means that 78% of the explained variance is accounted for by a single general factor. This means that the scale is probably essentially unidimensional and that its total unit-weighted score is reliable (however, some authors, e.g. Stucky & Edelen [2014], recommend values of .85 and higher so we should interpret even this conclusion with caution). On the other hand, as far as Compassion for the Self is concerned, the Omega hierarchical is .78, but the ECV value is only .53, which means that only 53% of explained variance is attributable to a single general factor, and the rest is explained by specific factors. The conclusion is that this scale is probably not essentially unidimensional, and its total unit-weighted score will be unreliable.

The patterns of the factor loading on the general factor are instructive in this respect (see Table S3 in Supplementary Materials): for the Compassion for Others scale, the factor loadings of eight items belonging to the Recognising Suffering and Understanding the Universality of Suffering subscales are on average .64, and the factor loadings of the twelve items belonging to Feeling for the Person Suffering, Tolerating Uncomfortable Feelings, and Acting or Being Motivated to Act to Alleviate Suffering scales are on average .68, which is a negligible difference. By contrast, for the Compassion for the Self scale, the factor loadings of the eight items belonging to the Recognising Suffering and Understanding the Universality of Suffering subscales are on average .29, whereas the factor loadings of the 12 items belonging to Feeling for the Person Suffering, Tolerating Uncomfortable feelings, and Acting or Being Motivated to act to Alleviate Suffering scales are on average .61 – this difference is huge and indicates that this scale is probably not unidimensional.

Discussion

The goal of this research study was to assess the psychometric properties and factor structure of the Slovak version of the newly developed Sussex-Oxford Compassion Scales (Gu et al., 2020). We agree with Strauss et al. (2016) that the lack of a common definition and statistically strong measuring instruments are restricting the research on compassion and self-compassion, as it is hard to measure them and to assess the effectiveness of interventions cultivating them. Our research results show that for the Sussex-Oxford Compassion for Others Scale (SOCS-O) the multidimensional definition of compassion for others fits the factor structure of the scale, but this does not apply to the Sussex-Oxford Compassion for the Self Scale (SCOS-S). Therefore, before we can develop and tailor a scale measuring it, self-compassion needs further analysis.

All self-compassions scales seem to have similar problems related to their total score. For example, the Self-Compassion Scale (SCS; Neff, 2003a) – the most frequently used instrument to measure self-compassion – does not show a good factor structure, and it is recommended that the total score of the scale should not be used. Instead, the score should be calculated separately for the positive and negative items of the SCS, which are Self-Compassionate Responding and Self-Uncompassionate Responding, as shown in samples from 11 different countries (Halamová et al., 2021). Similarly, the total score should not be used in relation to the CEAS scale (Gilbert et al, 2017) for compassion to self (CEAS-SC). Instead, the Engagement and Action of the CEAS-SC subscales should be calculated separately. Self-compassion is not unidimensional and has no strong single general factor in any of these SOCS-S, SCS, and CEAS-SC scales. The same appears to be true of the original data (Gilbert et al, 2017; Gu et al., 2020; Neff, 2003a), or at least that is what can be surmised from the limited information in the original articles. The conclusion then is that self-compassion has yet to be well-defined or operationalised as indicated by the problems using the total scores for the different self-compassion scales.

Our research findings cast doubt on whether a total score can be used with SOCS-S with either a Slovak sample or an English sample as the self-compassion of the SOCS-S does not show convincing enough support for an overarching general factor. However, SOCS-S and SOCS-O were designed as multidimensional measures. Use of the total score is recommended for SOCS-O, but the SOCS-S total score should not be used without further research on the interactions between the five subscales of self-compassion and self-compassion itself. According to our results regarding the SOCS-S, self-compassion may consist of two dominant factors, which could be called Rational Compassion (comprising two specific factors: Recognising Suffering and Understanding the Universality of Suffering), and Emotional/Behavioural Compassion (containing three specific factors: Feeling for the Person Suffering, Tolerating Uncomfortable Feelings, and Acting or Being Motivated to Act to Alleviate Suffering). It may be that Compassion can also be

divided into two dominant factors – Rational and Emotional/Behavioural Compassion – rather than a single general overarching factor – Compassion. These interpretations are based on the highly asymmetrical pattern of the magnitudes of factor loadings and asymmetry of factor intercorrelations. In future research, two-tier models should be fitted by treating the Rational and Emotional Compassion and the Behavioural Compassion factors as primary/dominant factors (the first tier) and the five-item factors as a set of secondary/specific factors (the second tier), see (Cai, 2010; Jeon et al., 2018 or Stucky & Edelen, 2014).

Given the asymmetry of the pattern of factor loading in our results as well as in Gu et al. (2020), we could seriously consider the hypothesis that there is no single general overarching factor (Compassion), but rather two dominant factors: Rational Compassion (encompassing two specific factors: Recognising Suffering and Understanding the Universality of Suffering), and Emotional/Behavioural Compassion (containing three specific factors: Feeling for the Person Suffering, Tolerating Uncomfortable Feelings, and Acting or Being Motivated to Act to Alleviate Suffering).

The SOCS was developed in parallel versions for compassion and self-compassion to enable close examination of the overlaps. Although the scale authors and other researchers do not distinguish between compassion and self-compassion (e.g. Gilbert, 2014; Gu et al., 2020), our research suggests there are some differences between them that need further exploration. The novelty and originality of our approach is that we interpreted our own data and the Gu et al. (2020) data in light of the bifactor modelling and we demonstrated that the assumption of essential unidimensionality (over and above the apparent multidimensionality) is valid for the Compassion for Others scale, but not for the Compassion for Self scale.

Recently, several academics have stated that self-compassion is a multidimensional construct (Gilbert et al, 2017; Jazaieri et al., 2013; Neff, 2003a; Strauss et al., 2016). However, they disagree over the dimensions of self-compassion. In future, more research using a qualitative methodology could shed more light on the nature of self-compassion, the dimensions and definition, and even the items that make up the measuring instrument to ensure it still represents the multidimensionality and that there is a strong general factor behind it. In addition, a qualitative research methodology could prove useful in identifying the differences and similarities between compassion and self-compassion (Baránková et al., 2019; Halamová et al., 2018) and provide insights into participant experience of the compassionate interventions and their impacts (Koróniová, Halamová, & Taňkošová., 2020).

Limitations and Future Research

Our limitations mainly stemmed from the non-randomised data gathering. Therefore, the extent to which the results can be generalised is quite limited. Given

that our research sample predominantly comprised young adult females, the limitations of this study also relate to gender and age. Nonetheless, the results regarding the relationships between compassion and self-compassion with age and gender are inconsistent (e.g. Bluth et al., 2017; Martins et al., 2013). In the original research by Gu et al. (2020), it was recommended that the scale should be translated into different languages to test cross-cultural differences, but there is always the risk of the translated instruments being understood differently and of cultural sensitivity. The third limitation is that we did not present and test more complex models (two-tier models) that assume the Rational and Emotional Compassion and the Behavioural Compassion factors are treated as primary/dominant factors (the first tier) and the five-item factors as a set of secondary/specific factors (the second tier); see Cai (2010), Jeon et al. (2018) or Stucky and Edelen (2014): this would have doubled the size of this study and so it will have to be considered in future research. The third limitation is that it would be interesting to test the measurement invariance of both instruments in relation to gender and age; however, the gender imbalance and relatively homogeneous age range of our sample restricts the potential for generalisation.

As the SOCS are self-reporting measurements, they are prone to social desirability and other sorts of bias. In future, it would be useful to analyse a triangulation of SOCS and non-self-reporting measures, such as heart rate variability (Halamová et al., 2019; Kirby et al., 2017), electromyography (Koróniová, Halamová, & Džongová, 2020), eye-tracking (Strnádelová et al., 2019), facial action units (Kanovský et al., 2020), and behavioural tasks (Gu et al., 2020) to diagnose level of compassion and self-compassion in people. In addition, this would be valuable for assessing the effectiveness of compassionate interventions.

Conclusion

The results of our research demonstrate that for the SCOS-O the multidimensional definition of compassion for others fits the factor structure of the scale, but this does not apply to the SCOS-S. Self-compassion has two dominant factors over and above five specific factors: Rational Compassion (containing two specific factors: Recognising Suffering and Understanding the Universality of Suffering) and Emotional/Behavioural Compassion (containing three specific factors: Feeling for the Person Suffering, Tolerating Uncomfortable Feelings, and Acting or Being Motivated to Act to Alleviate Suffering). Therefore, the total self-compassion score is unsuitable for use because the scale lacks essential unidimensionality; nonetheless, the compassion for others total score can be used safely.

References

- Baer, R. A., Carmody, J., & Hunsinger, M. (2012). Weekly change in mindfulness and perceived stress in a mindfulness-based stress reduction program. *Journal of Clinical Psychology, 68*, 755–765. <https://doi.org/10.1002/jclp.21865>
- Baránková, M., Halamová, J., & Koróniová, J. (2019). Non-expert views of compassion: Consensual qualitative research using focus groups. *Human Affairs, 29*(1), 6–19. <https://doi.org/10.1515/humaff-2019-0002>
- Benda J., & Vyhnanek, A. (2019, September 29–29). *Childhood, toxic shame, toxic guilt and self-compassion* [Paper presentation]. 7th International Conference on Pesso Boyden System Psychomotor® (PBSP®), Science and Good Practice, Prague, Czech Republic.
- Bluth K., Campo R. A., Futch W. S., & Gaylord S. A. (2017). Age and gender differences in the associations of self-compassion and emotional well-being in a large adolescent sample. *Journal of Youth and Adolescence, 46*, 840–853. <https://doi.org/10.007/s10964-016-0567-2>
- Cai, L. (2010). A two-tier full-information item factor analysis model with applications. *Psychometrika, 75*, 581–612. <https://doi.org/10.1007/s11336-010-9178-0>
- Cucina, J., & Byle, K. (2017). The bifactor model fits better than the higher-order model in more than 90% of comparisons for mental abilities test batteries. *Journal of Intelligence, 5*, 27. <https://doi.org/10.3390/jintelligence5030027>
- Davis, M. H. (1980). *A multidimensional approach to individual differences in empathy*. Retrieved from https://www.uv.es/~friasnav/Davis_1980.pdf
- Ekman, P. (2014). *Moving toward global compassion*. Paul Ekman Group.
- Gilbert, P. (2014). The origins and nature of compassion focused therapy. *British Journal of Clinical Psychology, 53*, 6–41. <https://doi.org/10.1111/bjc.12043>
- Gilbert, P., Catarino, F., Duarte, C., Matos, M., Kolts, R., Stubbs, J., Ceresatto, L., Duarte, J., Pinto-Gouveia, J., & Basran, J. (2017). The development of compassionate engagement and action scales for self and others. *Journal of Compassionate Health Care, 4*(1), 4. <https://doi.org/10.1186/s40639-017-0033-3>
- Goetz, J. L., Keltner, D., & Simon-Thomas, E. (2010). Compassion: An evolutionary analysis and empirical review. *Psychological Bulletin, 136*(3), 351–374. <https://doi.org/10.1037/a0018807>
- Gu, J., Baer, R., Cavanagh, K., Kuyken, W., & Strauss, C. (2020). Development and psychometric properties of the Sussex-Oxford compassion scales (SOCS). *Assessment, 27*(1), 3–20. <https://doi.org/10.1177/1073191119860911>
- Gu, J., Strauss, C., Crane, C., Barnhofer, T., Karl, A., Cavanagh, K., & Kuyken, W. (2016). Examining the factor structure of the 39-item and 15-item versions of the Five Facet Mindfulness Questionnaire before and after mindfulness based cognitive therapy for people with recurrent depression. *Psychological Assessment, 28*, 791–802. <https://doi.org/10.1037/pas0000263>

- Halamová, J., Baránková, M., Strnádelová, B., & Koróniová, J. (2018). Consensual qualitative research on free associations for compassion and self-compassion. *Human Affairs*, 28(3), 253–270. <https://doi.org/10.1515/humaff-2018-0021>
- Halamová, J., Kanovský, M., Petrocchi, N., Moreira, H., López, A., Barnett, M., Yang, E., Benda, J., Brähler, E., Zeng, X., & Zenger, M. (2021). Factor structure of the Self-Compassion scale in 11 international samples. *Measurement and Evaluation in Counseling and Development*, 54, 1–23. <https://doi.org/10.1080/07481756.2020.1735203>
- Halamová, J., Koróniová, J., Kanovský, M., Kénesy-Túnyiová, M., & Kupeli, N. (2019). Psychological and physiological effects of emotion focused training for self-compassion and self-protection. *Research in Psychotherapy: Psychopathology, Process and Outcome*, 22, 264–279. <https://doi.org/10.4081/ripppo.2019.358>
- Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, 44, 227–239. <https://doi.org/10.1348/014466505X29657>
- Hwang, J. Y., Plante, T., & Lackey, K. (2008). The development of the Santa Clara brief compassion scale: An abbreviation of Sprecher and Fehr's compassionate love scale. *Pastoral Psychology*, 56, 421–428. <https://doi.org/10.1007/s11089-008-0117-2>
- Jazaieri H., McGonigal K., Jinpa T., Doty, J. R., Gross, J. J., & Goldin, P. (2013). A randomized controlled trial of compassion cultivation training: Effects on mindfulness, affect, and emotion regulation. *Motivation and Emotion*, 38(1), 23–35. <https://doi.org/10.1007/s11031-013-9368-z>
- Jeon, M., Rijmen, F. & Rabe-Hesketh, S. (2018): CFA models with a general factor and multiple sets of secondary factors. *Psychometrika*, 83, 785–808. <https://doi.org/10.1007/s11336-018-9633-x>
- Kanovský, M., Baránková, M., Halamová, J., Strnádelová, B., & Koróniová, J. (2020). Analysis of facial expressions made while watching a video eliciting compassion. *Perceptual and Motor Skills*, 127(2), 317–346. <https://doi.org/10.1177/0031512519897512>
- Kirby, J. N., Doty, J. R., Petrocchi, N., & Gilbert, P. (2017). The current and future role of heart rate variability for assessing and training compassion. *Frontiers in Public Health*, 5, 40. <https://doi.org/10.3389/fpubh.2017.00040>
- Kline, P. (1999). *The handbook of psychological testing*. Routledge.
- Koróniová, J., Halamová, J., & Džongová, Z. (2020). Differences in electromyography during self-compassionate and self-critical imageries according to the level of self-criticism. *Sudia Psychologica*, 62(4), 364–375. <https://doi.org/10.31577/sp.2020.04.810>
- Koróniová, J., Halamová, J., & Taňkošová, N. (2020): Level of self-criticism and changes in imagery among participants attending Emotion Focused Training for Self-Compassion and Self-Protection. *Československá psychologie*, 64(5), 608–624.

- MacBeth, A., & Gumley, A. (2012). Exploring compassion: A meta-analysis of the association between self-compassion and psychopathology. *Clinical Psychology Review*, 32, 545–552. <https://doi.org/10.1016/j.cpr.2012.06.003>
- Martins D., Nicholas N.A., Shaheen M., Jones L., & Norris K. (2013). The development and evaluation of a Compassion Scale. *Journal of Health Care for the Poor and Underserved*, 24, 1235–1246. <https://doi.org/10.1353/hpu.2013.0148>
- Maslach, C., Jackson, S. E., & Leiter, M. P. (1981). *Maslach Burnout Inventory: MBI*. Consulting Psychologists press.
- McNeish, D. (2017). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, 23, 412–433. doi:10.1037/met0000144
- Muthén, L. K., & Muthén, B. O. (1998-2010). *Mplus User's Guide*. Sixth Edition. Muthén & Muthén.
- Neff, K. D. (2003a). Development and validation of a scale to measure self-compassion. *Self and Identity*, 2, 223–250. <https://doi.org/10.1080/15298860309027>
- Neff, K. D. (2003b) Self-compassion: An alternative conceptualization of a healthy attitude toward oneself. *Self and Identity*, 2, 85–102. <https://doi.org/10.1080/15298860309032>
- Raes, F., Pommier, E., Neff, K. D., & Van Gucht, D. (2011). Construction and factorial validation of a short form of the Self-compassion scale. *Clinical Psychology & Psychotherapy*, 18, 250–255. <https://doi.org/10.1002/cpp.702>
- Raykov, T., & Marcoulides, G. A. (2019). Thanks coefficient alpha, we still need you! *Educational and Psychological Measurement*, 79(1), 200–210. <https://doi.org/10.1177/0013164417725127>
- Reise, S. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667–696. <https://doi.org/10.1080/00273171.2012.715555>
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, 21, 137–150.
- Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika*, 75, 243. <https://doi.org/10.1007/s11336-009-9135-y>
- 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 Publishers.
- Savalei, V., & Reise, S. P. (2019). Don't forget the model in your model-based reliability coefficients: A reply to McNeish (2018). *Collabra: Psychology*, 5(1), Article 36. <https://doi.org/10.1525/collabra.247>
- Seppälä, E. M., Simon-Thomas, E., Brown, S. L., Worline, M. C., Cameron, C. D., & Doty, J. R. (2017). *The Oxford handbook of compassion science*. Oxford University Press.
- Soper, D. (2020). A-priori sample size calculator for structural equation models. <http://www.danielsoper.com>

- Sprecher, S., & Fehr, B. (2005). Compassionate love for close others and humanity. *Journal of Social and Personal Relationships*, 22(5), 629–651.
- Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences*. Routledge.
- Stewart-Brown, S., Tennant, A., Tennant, R., Platt, S., Parkinson, J., & Weich, S. (2009). Internal construct validity of the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS): A Rasch analysis using data from the Scottish Health Education Population Survey. *Health and Quality of Life Outcomes*, 7, 15. <https://doi.org/10.1186/1477-7525-7-15>
- Strauss, C., Taylor, B. L., Gu, J., Kuyken, W., Baer, R., Jones, F., & Cavanagh, K. (2016). What is compassion and how can we measure it? A review of definitions and measures. *Clinical Psychology Review*, 47, 15–27. <https://doi.org/10.1016/j.cpr.2016.05.004>
- Strnádelová, B., Halamová, J., & Kanovský, M. (2019). Eye-tracking of facial emotions in relation to self-criticism and self-reassurance. *Applied Artificial Intelligence*, 33(10), 839–862. <https://doi.org/10.1080/08839514.2019.1646004>
- Stucky, B. D., & Edelen, M. O. (2014). Using hierarchical IRT models to create unidimensional measures from multidimensional data. In S. P. Reise & D. A. Revicki (Eds.), *Handbook of item response theory modelling: Applications to typical performance assessment* (pp. 183–206). Routledge, Taylor & Francis Group.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- Williams, M. J., Dalgleish, T., Karl, A., & Kuyken, W. (2014). Examining the factor structures of the Five Facet Mindfulness Questionnaire and the Self-compassion scale. *Psychological Assessment*, 26(2), 407–418. <https://doi.org/10.1037/a0035566>

Faktorska struktura Sussex-oxfordske ljestvice suosjećanja

Sažetak

Kvalitetno istraživanje suosjećanja, za koje postoji interes u mnogim znanstvenim područjima, teško je provesti bez pouzdanoga i valjanog instrumenta suosjećanja i samosuosjećanja. Nedavno su razvijene Sussex-oxfordske ljestvice suosjećanja (SOCS; Gu i sur., 2020), s dvjema paralelnim formama za mjerenje suosjećanja (SCOS-O) i samosuosjećanja (SCOS-S) koje sadrže po 20 čestica. Budući da je riječ o novorazvijenim mjerama, provedenim je istraživanjem provjerena njihova faktorska struktura (Gu i sur., 2020). U istraživanje je uključeno 1080 sudionika, 19.35 % muškaraca i 80.65 % žena, prosječne dobi 29.29 godina ($SD = 11.15$). Korišten je prigodan uzorak, a sudionici su regrutirani putem društvenih mreža. Za obje su skale provjereni svi modeli koje navode autori (Gu i sur., 2020): jednofaktorski model, petofaktorski model i petofaktorski hijerarhijski model. Rezultati pokazuju da je za objašnjenje suosjećanja prema drugima (SCOS-O) prikladan multidimenzijski model, međutim, on nije dobar za objašnjenje samosuosjećanja (SCOS-S). Kod samosuosjećanja moguće je izdvojiti dva nadređena faktora: Racionalno suosjećanje (koje sadrži dva specifična faktora: Prepoznavanje patnje i Razumijevanje univerzalnosti patnje) i Emocionalno/ponašajno suosjećanje (koje sadrži tri specifična faktora: Osjećaj za osobu koja pati, Toleriranje neugodnih osjećaja i Djelovanje ili motiviranost za djelovanje da bi se ublažilo patnju).

Stoga kod ljestvice samosuosjećanja nije prikladno koristiti ukupan rezultat jer nije riječ o jednodimenzijskome konstrukt, dok se kod ljestvice suosjećanja prema drugima ukupan rezultat smije koristiti.

Ključne riječi: suosjećanje, samosuosjećanje, psihometrijska analiza, faktorska analiza, procjena

Primljeno: 12. 10. 2020.