Original Research Article

Assessment
1–16

DOI: 10.1177/10731911221137540
journals.sagepub.com/home/asm
SAGE

Self-Compassion in Chinese Young Adults: Its Measurement and Measurement Construct

Mengya Zhao[1], Tamsin Ford[2], Janet Smithson[1], Peng Wang[3], and Anke Karl[1]

Abstract

Objectives. Self-compassion is the ability to be kind to oneself in adversity. This multidimensional construct is typically assessed by the Self-Compassion Scale (SCS). In Chinese samples, there have been inconsistent psychometric findings that impede cross-cultural research. This study aimed to explore the factor structure of the Chinese version (SCS-C). *Methods.* Two samples of young Chinese adults were recruited (Sample 1, N = 465, 141 men, Mean age [M_{age}] = 20.26; Sample 2, N = 392, 71 men; M_{age} = 18.97). Confirmatory factor analyses and exploratory structural equation modeling (ESEM) were used to examine previously reported four- and six-factor structures of SCS-C. *Results.* Although ESEM supported the six-factor structure when a problematic item was omitted, we found stronger evidence for a novel four-factor structure of the SCS-C revealed with self-kindness, common humanity, mindfulness, and uncompassionate self-responding. This suggests that Chinese individuals have a different understanding of the negative components of the original self-compassion definition, which was based on the United States and other mostly Western samples. Omega coefficients of the bifactor models suggested that using the SCS total score in Chinese samples is inappropriate. However, high factor determinacy and construct replicability indicated that the general factor of SCS-C could be used in a structural equation modeling context for both four-factor and six-factor structures. *Conclusions.* When using the existing SCS-C in path models, researchers should use a latent variable approach and establish the measurement construct rather than sum scores of the scale or subscales without checking the factor structure in future empirical studies. Also, the SCS-C needs to be revised, and we proposed directions forward for future research.

Keywords

self-compassion scale, Chinese version, factor analysis, exploratory structural equation modeling, bifactor model

Self-compassion means being kind to oneself when facing difficulties (Neff, 2003). According to the predominant definition, self-compassion is a multidimensional construct about how we relate to ourselves when facing challenges or our inadequacies (Neff, 2003, 2016) Individuals can respond to challenges or personal inadequacies with compassion, kindness and understanding (self-kindness) rather than harshly judging or blaming themselves (self-judgment). In addition, individuals may understand their difficulties, failure, or inadequacies as a shared experience with other human beings (common humanity) rather than feeling the difficulties only happened to themselves (isolation). Furthermore, to process a difficult situation, individuals use a mindful and balanced way (mindfulness) rather than being carried away by their emotions (over-identification).

Although the concept of self-compassion is rooted in Ancient Buddhist tradition (Neff, 2003), only in the last two decades has predominantly Western academic psychology explored its role as a protective factor for mental health. This research has demonstrated a positive association between self-compassion and mental well-being in adolescents (Marsh et al., 2018) and adults (Zessin et al., 2015). Self-compassion can be cultivated using psychological interventions for healthy individuals and clinical patients (Germer & Neff, 2013). In addition, self-compassion has been positively associated with social functioning, such as increased interpersonal trust and social support provision (Crocker & Canevello, 2008) and better relationship conflict resolutions (Yarnell & Neff, 2013).

Research into self-compassion has increased in China over the last decade. For example, cross-cultural studies revealed higher levels of self-compassion in American

[1]University of Exeter, Exeter, UK
[2]University of Cambridge, Cambridge, UK
[3]Shandong Normal University, Jinan, China

Corresponding Author:
Mengya Zhao, Psychology, University of Exeter, Washington Singer Laboratories, Perry Road, Exeter EX4 4QG, UK.
Email: m.zhao@exeter.ac.uk

compared with participants from Taiwan (Neff et al., 2008) and larger negative associations between self-compassion and depressive symptoms in the United Kingdom compared with Chinese adolescents (Zhao, Ford et al., 2021). Other research has replicated positive associations between self-compassion and mental health, such as less nonsuicidal self-harm in Chinese adolescents (Jiang, You, Ren et al., 2017; Jiang, You, Zheng et al., 2017), higher life satisfaction (Y. Yang et al., 2016), lower depressive symptom and fewer conflicts in friendships in youth (Zhao, Ford et al., 2021) and higher relationship harmony in young adults (X. Yang, 2016). Although this research substantially contributed to the understanding of self-compassion in the Chinese culture, emerging psychometric findings (Neff et al., 2019; Tsai, 2015; Zeng et al., 2016) raise some issues about the measurement and potentially the cross-cultural conceptualization of the construct (Zhao, Smithson et al., 2021).

To date, the Self-Compassion Scale (SCS; Neff, 2003) has been the most widely used measure in research related to self-compassion and has been translated into nineteen different languages (Neff et al., 2019). The SCS consists of 26 items on a 5-point Likert-type scale (from "almost never" to "almost always") and has six subscales: self-kindness, self-judgment, common humanity, isolation, mindfulness, and over-identification (Neff, 2003; Neff et al., 2019). The scale was originally designed to assess self-compassion as a total score or separate scores of the six subscales (Neff et al., 2019). However, there has been a lively scientific debate about factor structure and scoring of the SCS.

For the original English version and various translations, confirmatory factor analysis (CFA) studies supported the proposed six-factor structure where the subscales described above are intercorrelated (the six-factor correlated model; Model 3 in Figure 1) or load onto a single factor self-compassion (the higher-order model) (Neff et al., 2019). Hence, scoring suggestions for the SCS as one total score or six sub-scores have been the most widely used to date (Neff et al., 2019). Alternative research has proposed two separate scores for the SCS with its three positive (self-kindness, common humanity, and mindfulness) and three negative (self-judgment, isolation, and overidentification) subscales rather than one total score (Model 5 in Figure 1; Brenner et al., 2017; Halamová et al., 2020; López et al., 2015). This way of scoring the SCS has been theoretically associated with the social mentalities theory, which postulates that in addition to a safeness system (self-compassion), there is a separate threat-defense system that, when activated, leads us to treat ourselves with self-criticism (Brenner et al., 2017; Gilbert, 2005). Psychometric support has been obtained from several studies for a factor "compassionate self-responding" (including self-kindness, common humanity, and mindfulness) and a factor "uncompassionate self-responding" (including self-judgment, isolation, and overidentification) using the original and translated versions of the SCS (Brenner et al., 2017; Costa et al., 2016; Halamová et al., 2020; López et al., 2015). Furthermore, some research did not find support for scoring SCS as a total score (e.g., Williams et al., 2014). In addition, some suggest that the differential association of compassionate and uncompassionate self-responding with mental health outcomes further supports a two-factor structure which means scoring SCS as two separate scores (e.g., Brenner et al., 2018; Körner et al., 2015). However, others failed to confirm this proposition (Neff et al., 2019). Meanwhile, some researchers suggest only using the subscales representing compassionate self-responding (Muris & Otgaar, 2020).

Psychometric research has sought to address the controversy about the structure and scoring of SCS by comparing various models using different statistical approaches (e.g., Neff et al., 2019). Further to the original higher-order model, a bifactor model was tested and deemed the more appropriate approach because all latent variables (self-compassion and its components) are directly defined by the observed variables (specific items; Neff et al., 2017, 2019; Rodriguez et al., 2016b). In contrast, the strict assumptions of higher-order models only allow an indirect hierarchical structure where specific items define subscales/factors and these, in turn, define the full scale/global factor (Gignac, 2016). Neff (2016) argued that the direct hierarchical structure inherent to bifactor models is in line with the definition of self-compassion and hence more suitable when exploring the scoring of the SCS. Despite this advantage, the bifactor model did not fully address issues raised in the debate. When exploring the SCS factor structure in clinical samples, the model fits of the bifactor model were not acceptable (Neff et al., 2017). Therefore, exploratory structural equation modeling (ESEM) was explored, which allows items to cross-load across different factors, whereas CFA only allows the items to load on one factor (Asparouhov & Muthén, 2009). Neff et al. (2019) argued that this is warranted because self-compassion is a multidimensional system of associated components (Neff, 2016), and ESEM is specifically designed for models considered a multidimensional system of different factors (Morin et al., 2013). Emerging research using ESEM explored the issue of SCS factor structure and scoring and found good model fits across different samples using the original English version of the SCS (7 samples) and several translated versions (13 samples) (Neff et al., 2019). Comparing five different models, Neff et al. (2019) found that ESEM consistently outperformed CFA in terms of model fits and supported the use of the SCS's original six subscale scores or a total score, whereas no evidence for two separate scores of compassionate self-responding and uncompassionate self-responding was found. In addition, the authors recommended that ESEM should be used for determining the factor structure of the SCS (Neff et al., 2019).

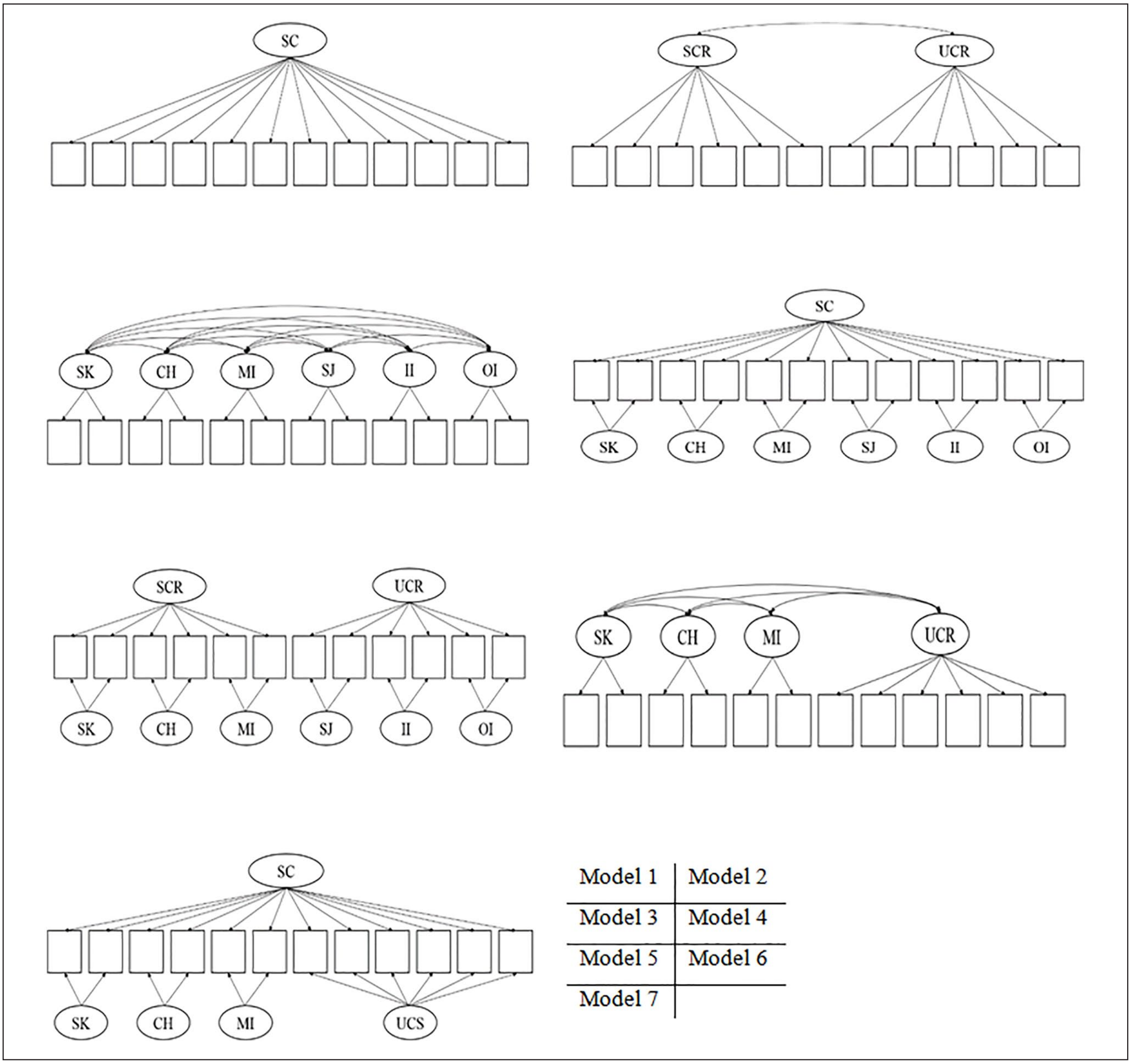

Figure 1. Diagrams of the Models

Note. SC = self-compassion; SCR = compassionate self-responding (including self-kindness, common humanity, and mindfulness); UCR = uncompassionate self-responding (including self-judgement, isolation, and overidentification); SK = self-kindness; CH = common humanity; MI = mindfulness; SJ = self-judgment; II = isolation; OI = overidentification.

Despite the use of ESEM, Neff et al. (2019) had model identification problems with the Chinese version of the SCS (SCS-C). In addition, existing psychometric evaluations were based on different translations of the SCS (SCS-C; see Table S1 in Supplementary Materials), with mixed support for construct validity. Interestingly, four previous studies did not satisfactorily replicate the higher-order six-factor model; Deng et al., 2012; Tsai, 2015; Zeng et al., 2016), and there is growing evidence for a different, not previously observed factor structure of the SCS in Chinese samples. After failing to replicate both the six-factor correlated model and its higher-order model, Tsai (2015) proposed a novel four-factor structure comprising self-kindness, common humanity, mindfulness, and one factor on which all items from negative subscales self-judgment, isolation, and overidentification loaded (which corresponds with uncompassionate self-responding; Model 6 in Figure 1). Similarly, Zeng et al. (2016) found evidence for three correlated positive but only one negative factor, comprising three highly correlated negative components, in community samples who identified as Buddhist or non-religious. These two studies (Tsai, 2015; Zeng et al., 2016) raise a particular problem with specifying the hypothesized three distinct negative components of the original self-compassion definition in the Chinese sample. These findings suggest that further research on the understanding of self-compassion in Chinese samples should examine a four-factor structure in addition to the original conceptualization. Unfortunately, Chinese validation studies did not further discuss why the novel factor structures may have emerged (Tsai, 2015; Zeng et al., 2016).

There are several possibilities to explain the discrepancy between English and Chinese versions of the SCS. The first could lie in the statistical approaches used. The Chinese validation studies used CFA; Deng et al., 2012; Jing et al., 2011; Tsai, 2015; Zeng et al., 2016), whereas recent research recommended ESEM to account for the synergetic factor structure (Neff et al., 2019). However, there were other problems with the SCS-C that further challenged the six-factor structure in Chinese samples. Across studies, low internal consistency of certain subscales was observed. In undergraduate samples, this was found for the self-judgment subscale, with Cronbach's alphas of .64 and .51 (Jing et al., 2011). Only the mindfulness subscale ($\alpha = .70$) in and the self-kindness subscale ($\alpha = .74$) in Jing et al. (2011) revealed acceptable internal consistencies, whereas, for the whole scale, it was good, with $\alpha = .84$ and .87 (Jing et al., 2011) despite the fact that the one higher-order factor structure was not supported in these studies. In Deng et al.'s (2012) study of an adolescent sample, most subscales ($\alpha =$ .58–.68) and even the full scale ($\alpha = .64$) had low internal consistency. This accumulation of psychometric problems revealed for the SCS-C raises the possibility that cultural differences in the understanding of the scale could be at the root of these measurement difficulties (see Zhao, Smithson et al., 2021 for detailed discussion). However, the statistical issues need to be addressed first.

In summary, research regarding psychometric properties of the SCS-C has revealed several limitations or gaps. First, different studies support different factor models, and the factor structure of SCS-C requires further discussion and replication. Second, some subscales have low internal consistency. Previous studies did not provide specific suggestions on how to use the scale in light of the potential problems with factor structure and internal consistency. More importantly, there are several different versions of the SCS-C, and studies of these measures have not used the bifactor model and ESEM as recommended to study the factor structure and generate scoring recommendations. Research that addresses the gaps in our understanding of the factor construct of SCS-C in a Chinese population of young adults is, therefore, necessary.

The current study aims to test the factor structure of the SCS-C using two samples and provide a recommendation for the use of SCS-C. For this, the study intents to test seven potential factor models (one-factor, two-factor correlated, six-factor correlated, single-bifactor, and two-bifactor models) following Neff et al. (2019) and another novel four-factor structure (four-factor correlated and its single-bifactor) following Tsai (2015) and Zeng et al.'s (2016) work.

Method

Participants and Procedure

We recruited two separate samples for this study. Sample 1 was recruited via social media, and participants in Sample 2 completed the survey in a paper-and-pen format. We used both samples to explore the factor structure. The inclusion criteria were being aged over 18 years, native Chinese speaker and university student (undergraduates or postgraduates). The study was approved by the Ethics Committee of the University (United Kingdom).

Recruitment of two samples was necessary because we initially aimed to test validity (including construct, convergent, and discriminant validity) following the measurements and procedure conducted by Neff (2003) and Neff et al. (2019). Moreover, two samples allowed us to examine convergent and discriminant validity separately and to reduce the burden on our participants.

Sample 1. Participants comprised 465 Chinese university students (141 men and 324 women, $M_{age} = 20.26$, $SD_{age} =$ 2.18). Data collection was conducted via an online platform (https://www.wjx.cn/), which has been widely used for survey-based research in China. All participants were recruited by social networks (e.g., WeChat). The participant information sheet and debriefing sheet were provided online. Participants gave informed consent and received 5 yuan as a reimbursement for their time.

Sample 2. A convenience sample of participants was recruited from a university in Northern China. Participants included 392 undergraduate students (71 men; 317 women. $M_{age} = 18.97$; $SD_{age} = .98$). All participants were given the participant information sheet, consent form, and debriefing sheet. After finishing the questionnaire, two gel pens were given as a reimbursement for their time.

Measures

Self-Compassion Scale Chinese Version (Both Samples). The SCS (Neff, 2003) includes 26 items using a 5-point Likert-type scale (from 1 "almost never" to 5 "almost always"). The Chinese version of the scale was based on the version from You's research team (Jiang, You, Ren et al., 2017; Jiang, You, Zheng et al., 2017). Two postgraduates (YL and KH) who majored in translation revised several words and re-translated several items according to explanations provided by a native English speaker (AC) who was a PhD researcher in Psychology. The translated version was reviewed by the first author (MZ) and then back-translated into English by two postgraduate students (JL and MW) who are bilingual (Chinese is their native language) and were studying psychology in English-speaking countries to ensure the language relevance and cultural correctness. The co-author (AK), a clinical psychologist and researcher into self-compassion, reviewed two back-translated versions. A mutual agreement between the first author (MZ) and the co-author (AK) was reached on the final SCS-C. This self-compassion scale Chinese version was used in both samples.

Data Analysis Approach

SPSS 23.0, Mplus 7.0, and R were used to analyze data. All factor analyses were conducted using Mplus syntax provided by Neff et al. (2019), apart from the novel four-factor structure models, which were not tested in Neff's study. Weighted least squares mean and variance adjusted estimation was used for all analyses, and pairwise deletion as the default way to deal with missing data was applied. Pairwise deletion was conducted due to a low percentage of missing data and missingness at completely random. Specifically, Sample 1 has no missing data across 26 items; Sample 2 has a very low percentage of missing data for individual items (the highest percentage was 2.8%), and 89% of participants filled all 26 items. Little's MCAR test demonstrated that the missingness is completely at random, $\chi^2 = 660.51$, $df = 664$, $p = .53$.

Seven correlated and bifactor CFA and ESEM models were tested (see Figure 1 for the model diagrams); Model 1 tested a one-factor model; Model 2, a two-factor correlated model; Model 3, a six-factor correlated model; Model 4, a single-bifactor model with a general factor and six specific factors; Model 5: a two-bifactor model including two general factors each with three specific factors; Model 6, a four-factor correlated model; and Model 7, a single-bifactor model with a general factor and four specific factors. In model assessments, we assessed the fit of models using the following benchmarks: the comparative fit index (CFI; $>.95$ for good, $>.90$ for acceptable), the Tucker–Lewis index (TLI; $>.95$ for good, $>.90$ for acceptable), the root mean square error of approximation (RMSEA; $<.06$ for good, $<.08$ for acceptable) with its 90% confidence interval (CI; Hu & Bentler, 1999; Marsh et al., 2005). Also, we also presented parameter estimates (e.g., factor loadings and inter-factor correlations) to evaluate the measurement models.

For the models that investigated the bifactor structure, we used the R package "Bifactor Indices Calculator" (Dueber, 2020) to evaluate omega coefficients to assess if the total score of the SCS-C reflects variations of the general factor in the bifactor models, in other words, if the SCS-C represents a general construct self-compassion. Following the suggestions by Rodriguez et al. (2016b), we determined coefficients omega (omega), omega hierarchical (omegaH), factor determinacy (FD), and construct reliability (H). Omega reflects the proportion of variance in the unit-weighted total score, which can be attributed to all sources of common variance (Reise, Bonifay et al., 2013; Revelle & Zinbarg, 2009), with scores of above .60 indicating acceptable and above .70 indicating good variance attribution (Bagozzi & Yi, 1988), as used in previous research (Brenner et al., 2017; Neff et al., 2019). OmegaH reflects the proportion of variance in total scores attributable to a single general factor (Reise, Moore et al., 2013). To evaluate omega coefficients, Rodriguez et al. (2016b) proposed to determine the ratio of omegaH/omega as the percentage of the reliable variance in the total score that can be attributed to the general factor. Recent research (Reise, Bonifay et al., 2013) suggested 75% as a benchmark for the ratio of omegaH/omega, which was applied as criterion in Neff et al. (2019). In addition, we used FD and H to assess if the general factor that was generated by the items is trustworthy and appropriate to be used in structural equation modeling (SEM) contexts. The factor score estimates are trustworthy when FD is $>.90$ (Gorsuch, 1983). H reflects if the latent variable is defined well by the relevant items, which is the case when H is $>.70$ (Hancock & Mueller, 2001).

Results

Testing Model Fits and Factor Loadings

Full Version of SCS-C. Model fit indices for our hypothesized factor models (Model 1–Model 7) are shown in Table 1. All models using CFA did not provide an adequate fit to the data. Using ESEM in Sample 1, we found the acceptable model fits for six-factor correlated model (Model 3), four-factor correlated model (Model 6), and single-bifactor model with four specific factors (Model 7). In Sample 2, the acceptable model fits were revealed for four-factor correlated model (Model 6). Consistent support across both samples was observed for Model 6 based on the model fits.

We then examined the standardized item factor loadings for the ESEM solutions with acceptable model fits, presented in Table S2, Table2, and Table S3. We found cross-loadings for some items when examining the six-factor correlated model (Model 3) in Sample 1 (Table S2). Also, we identified three items that did not significantly load on the factors as hypothesized: for self-judgment, Item 1 ("I'm disapproving and judgmental about my own flaws and inadequacies"), and for mindfulness, Item 9 ("When something upsets me I try to keep my emotions in balance") and Item 22 ("When I'm feeling down I try to approach my feelings with curiosity and openness"). Overall, the six factors were acceptably defined ($\lambda = .11$ to $.80$, $M\lambda = .51$). As for the

Table 1. Model Fit Indices for the Alternative Models Tested (SCS-C).

Model	Sample	CFA χ^2	CFA *df*	CFA CFI	CFA TLI	CFA RMSEA [90% C.I.]	CFA SRMR	ESEM χ^2	ESEM *df*	ESEM CFI	ESEM TLI	ESEM RMSEA [90% C.I.]	ESEM SRMR
Full SCS-C													
Model 1	1	1,993.61***	299	.72	.690	.11 [.11, .12]	.09	1,993.67***	299	.72	.69	.11 [.11, .12]	.09
	2	2,028.26***	299	.66	.628	.12 [.12, .13]	.10	2,028.26***	299	.66	.62	.12 [.12, .13]	.10
Model 2	1	1,363.31***	298	.82	.805	.09 [.08, .09]	.07	1,224.63***	274	.84	.81	.09 [.08, .09]	.06
	2	1,269.59***	298	.81	.790	.09 [.09, .10]	.08	991.81***	274	.86	.83	.08 [.08, .09]	.06
Model 3	1	1,214.22***	284	.84	.821	.08 [.08, .09]	.06	**382.08*****	**184**	**.97**	**.94**	**.05 [.04, .06]**	**.03**
	2	1,000.46***[a]	284	.86	.838	.08 [.08, .09]	.07	399.12***[a]	184	.96	.93	.06 [.05, .06]	.03
Model 4	1	1,076.77***[a]	273	.87	.839	.08 [.08, .09]	.07	314.13***[a]	164	.98	.95	.04 [.04, .05]	.02
	2	1,204.90***[a]	273	.82	.781	.09 [.09, .10]	.08	321.42***[a]	164	.97	.94	.05 [.04, .06]	.03
Model 5	1	641.09***[a]	272	.94	.926	.05 [.05, .06]	.05	251.32***[a]	157	.98	.97	.04 [.03, .04]	.02
	2	759.87***[a]	272	.90	.885	.07 [.06, .07]	.06	271.37***[a]	157	.98	.95	.04 [.03, .05]	.02
Model 6	1	1,264.40***	293	.84	.819	.08 [.08, .09]	.06	**609.09*****	**227**	**.94**	**.91**	**.06 [.05, .07]**	**.04**
	2	1,134.52***	293	.83	.815	.09 [.08, .09]	.07	**566.96*****	**227**	**.93**	**.90**	**.06 [.06, .07]**	**.04**
Model 7	1[b]							**467.50*****	**205**	**.96**	**.93**	**.05 [.05, .06]**	**.03**
	2[b]							471.54***[a]	205	.95	.92	.06 [.05, .06]	.03
SCS-C (without 21)													
Model 1	1	1,511.56***	275	.78	.76	.10 [.09, .10]	.08	1,511.56***	275	.78	.76	.10 [.09, .10]	.08
	2	1,678.39***	275	.71	.69	.11 [.11, .12]	.09	1,678.39***	275	.71	.69	.11 [.11, .12]	.09
Model 2	1	871.95***	274	.90	.89	.07 [.06, .07]	.06	759.99***	251	.91	.89	.07 [.06, .07]	.05
	2	973.40***	274	.86	.84	.08 [.08, .09]	.07	726.33***	251	.90	.88	.07 [.06, .08]	.05
Model 3	1	726.57***[a]	260	.91	.91	.06 [.06, .07]	.05	302.02***[a]	165	.98	.96	.04 [.04, .05]	.03
	2	754.93***[a]	260	.90	.88	.07 [.06, .08]	.06	**318.03*****	**165**	**.97**	**.94**	**.05 [.04, .06]**	**.03**
Model 4	1	1,048.88***[a]	250	.86	.83	.08 [.08, .09]	.07	246.70***[a]	146	.98	.96	.04 [.03, .05]	.02
	2	1,171.83***	250	.81	.77	.10 [.09, .10]	.08	**271.76*****	**146**	**.97**	**.95**	**.07 [.04, .06]**	**.02**
Model 5	1	610.03***[a]	249	.94	.92	.06 [.05, .06]	.05	**205.73*****	**139**	**.99**	**.98**	**.03 [.02, .04]**	**.02**
	2[b]							**237.44*****	**139**	**.98**	**.96**	**.04 [.03, .05]**	**.02**
Model 6	**1**	**762.89*****	**269**	**.91**	**.90**	**.06 [.06, .07]**	**.06**	**479.92*****	**206**	**.95**	**.93**	**.05 [.05, .06]**	**.03**
	2	819.35***	269	.89	.87	.07 [.07, .08]	.06	**548.46*****	**206**	**.95**	**.93**	**.06 [.05, .06]**	**.04**
Model 7	1	626.88***[a]	250	.93	.92	.06 [.05, .06]	.05	**369.50*****	**185**	**.97**	**.95**	**.05 [.04, .05]**	**.03**
	2	**649.07*****	**250**	**.92**	**.90**	**.06 [.06, .07]**	**.05**	**391.18*****	**185**	**.96**	**.93**	**.05 [.05, .06]**	**.03**

Note. Bold values mean acceptable model fits. These solutions had model identification issues, suggesting over parameterization. Model 1, one-factor model; Model 2, two-factor correlated model; Model 3, six-factor correlated model; Model 4 single-bifactor model with six factors and one general factor; Model 5, two-bifactor model with six factors and two general factor; Model 6, four-factor correlated model; Model 7, single-bifactor model with four4 factors and one general factor. SCS = self-compassion scale; CFA = confirmatory factor analysis; ESEM = exploratory structural equation modeling; CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; CI = confidence interval.

[a]Model has identification issues with Mplus warning message. [b] Mplus warning message, the residual covariance matric (theta) is not positive definite, not model fit results.

Table 2. Standardized Factor Loadings for Model 6 in Sample 1 and Sample 2 (Full SCS-C).

	Sample 1				Sample 2			
Items	SK(λ)	CH(λ)	MI(λ)	USR(λ)	SK(λ)	CH(λ)	MI(λ)	USR(λ)
SK								
SC5	**.60*** **	.17***	−.06	−.16***	**.64*** **	.02	.17***	−.07
SC12	**.66*** **	.12*	.13***	−.06	**.84*** **	.07	.11**	−.09**
SC19	**.58*** **	.14**	.01	.03	**.64*** **	.08	.10**	.003
SC23	**.56*** **	−.22***	−.07	.24***	**.32*** **	.05	−.12**	.32***
SC26	**.47*** **	.06	−.23***	.25***	**.37*** **	.07	.10*	.31***
CH								
SC3	.10	**.41*** **	−.16**	.15**	.27***	**.22*** **	.07	.01
SC7	.01	**.72*** **	.11**	−.04	−.01	**.67*** **	.03	−.01
SC10	.04	**.54*** **	.08*	.02	−.003	**.70*** **	.01	−.04
SC15	.14*	**.44*** **	−.04	.08	−.02	**.57*** **	.14**	.07
MI								
SC9	.28***	.33***	**−.07**	.12**	.36	.09	**.43*** **	.12**
SC14	.42***	.28***	**−.06**	.14**	.17**	.30***	**.24*** **	.18***
SC17	.38***	.19***	**−.19*** **	.15***	.34***	.17***	**.27*** **	.10*
SC22	.26***	.27***	**.02**	.13**	.14**	.25**	**.36*** **	.19***
USR								
SC1	.01	−.03	−.05	**.41*** **	.05	−.05	−.02	**.44*** **
SC8	.15**	−.14*	.72***	**.18***	−.05	.13*	−.63***	**.46*** **
SC11	.38***	−.28***	−.10*	**.36*** **	.35***	−.12*	.09	**.42*** **
SC16	.07	−.21***	.02	**.55*** **	−.05	−.26***	.03	**.65*** **
SC21	−.01	−.08	.82***	**.40*** **	.15**	.04	−.66***	**.49*** **
SC4	.08	−.04	−.09*	**.58*** **	−.14**	−.10*	.18***	**.68*** **
SC13	−.04	.07	−.08	**.67*** **	.04	−.08	.15***	**.57*** **
SC18	−.02	−.03	−.08*	**.63*** **	.15**	.01	.05	**.51*** **
SC25	−.02	.11**	−.07	**.75*** **	.09	.09*	.10*	**.68*** **
SC2	−.19**	.17**	.003	**.62*** **	−.25***	.14**	.22***	**.63*** **
SC6	−.04	−.05	−.09*	**.68*** **	−.13**	.03	.13**	**.69*** **
SC20	−.14**	.05	−.08	**.55*** **	.05	−.07	.15***	**.52*** **
SC24	−.10*	.13	−.05	**.59*** **	−.09	.08	.18***	**.66*** **

Note. Target factor loadings are in bold. λ = standardized factor loadings. SCS = self-compassion scale; SK = self-kindness; CH = common humanity; MI = mindfulness; USR = uncompassionate self-responding.
*$p < .05$. ** $p < .01$. ***$p < .001$.

correlation between factors, self-judgment was not significantly correlated with common humanity ($r = −.05$) and mindfulness ($r = .03$), and overidentification was not significantly correlated with common humanity ($r = −.02$). The correlation between self-kindness, common humanity, and self-judgment was significant, $r = .11$ to .45, $M_r = .27$.

For the four-factor correlated model (Model 6; Table 2), we found in Sample 1 that mindfulness was poorly defined with either significantly negative factor loading (Item 17: "When I fail at something important to me, I try to keep things in perspective") or no significant loadings. The other three factors were overall acceptably defined (λ = .18 to .72, $M\lambda = .53$). There were cross-loadings for several items. As for the correlation between factors, mindfulness was not significantly correlated with self-kindness ($r = −.04$), and negatively correlated with common humanity ($r = −.16$, $p < .001$) and uncompassionate self-responding ($r = −.11$, $p < .001$). In Sample 2, the factor loadings of the items were significant, and the four factors were acceptably defined (λ = .22 to .84, $M\lambda = .53$), but several items had cross-loadings. The correlation between the four factors, self-kindness, common humanity, mindfulness and uncompassionate self-responding, was significant ($r = .15$ to.43, $M_r = .25$).

For the single-bifactor model with four specific factors in Sample 1 (Table S3), the factor loadings of the general factor, although significant, were fairly defined ($|\lambda| = .18$–.69, $M\lambda = .45$), and there are two items with small negative factor loading, Items 8 ("When times are really difficult, I tend to be tough on myself") and 21 ("I can be a bit cold-hearted toward myself when I'm experiencing suffering respectively"). As for the specific factors, similar to the

Table 3. Potential Problematic Item/Factor According to Mplus Warning Message.

	Sample 1		Sample 2	
Models	CFA	ESEM	CFA	ESEM
Full SCS-C				
Model 1	—	—	—	—
Model 2	—	—	—	—
Model 3	—	—	OI	SC21
Model 4	SC14	SC10	SC21	SC21
Model 5	SC17	SC10	SC14	SC21
Model 6				
Model 7	SC17		SC17	SC21
SCS-C no 21				
Model 1				
Model 2				
Model 3	Overidentification	SC10	Mindfulness	
Model 4	SC14	SC10		
Model 5	SC17		SC14/SC24	
Model 6				
Model 7	SC17			

Note. SCS = self-compassion scale; CFA = confirmatory factor analysis; ESEM = exploratory structural equation modeling.

results of the four-factor correlated model, the factors were fairly defined (λ = .22 to .82, $M\lambda$ = .39 apart from mindfulness which was not defined by significant loadings.

In summary, although ESEM revealed acceptable solutions in Sample 1 and/or 2, all models had some factor loading problems, which included some non-significant loadings or low factor loadings, as well as cross-loadings of some items. Also, although the four-factor correlated model (Model 6) was supported with an acceptable model fit in both samples, the mindfulness subscale was not defined in Sample 1.

The Version of SCS-C Without Item 21. Based on the Mplus warning message, there were four potential items, which could lead to model identification problems (Table 3): Item 21, Item 14, Item 10, and Item 17. We decided to delete Item 21 for the following reasons: first, Item 21 was reported to be problematic in many models. Second, among the ESEM models with acceptable fits but with model identification problems (Table 1, Model 3–Model 5), there were two potential problematic items; Item 10 in Sample 1 and Item 21 in Sample 2. Furthermore, we checked the inter-item correlation (polychoric correlation) for all problematic items stated above, and we used the cut-off point .20 to .40 (Piedmont, 2014). In Sample 1, the inter-item correlations between Item 21 and other 24 items were below .20 and with only one item above .40; in Sample 2, the correlations between Item 21 and other 15 items were below .20 and with 1 item above .40.

The results of the ESEM for the SCS-C without Item 21 are shown in Table 1. The model solution using ESEM performed better than using CFA, but the four-factor correlated model (Model 6) in Sample 1 and the single-bifactor model with four specific factors (Model 7) in Sample 2 using CFA were supported with acceptable model fits. The two-bifactor model with six specific factors (Model 5), the four-factor correlated model (Model 6) and the single-bifactor model with four specific factors (Model 7) had acceptable model fits in both samples using ESEM. However, the six-factor correlated model (Model 3) using ESEM was only supported in Sample 2; in Sample 1, there was an over-identification issue. The bifactor solution, the one bifactor model with six specific factors (Model 4), was only supported in Sample 2.

The findings related to the factor loadings can be found in the Supplementary Material. Although the two-bifactor model with six specific factors (Model 5) had good model fits, the general factors were poorly defined in both samples (Table 4). In Sample 1, only 4 out of 13 items significantly loaded on compassionate self-responding and 2 out of 12 items on uncompassionate self-responding. Similarly, in Sample 2, there were only four items significantly loading on compassionate self-responding and one item on uncompassionate self-responding (Table 4). For the four-factor correlated model (Model 6), the findings resembled that of the full version. In Sample 1, mindfulness was not defined well, nor was it significantly correlated with any other factors. In Sample 2, the four factors were defined well, and the factors were intercorrelated with each other (Table S4). For the single-bifactor model with four specific factors (Model 7; Table S5), the general factor was fairly defined in both samples (Sample 1, λ = .30 to .70, $M\lambda$ = .45; Sample 2, λ = .15–.85, $M\lambda$ = .43), but Item 8 was not loaded on the

Table 4. Standardized Factor Loadings for the Model 5 Solution Using ESEM in Sample 1 and Sample 2 (SCS-C Omitting Item 21).

	Sample 1								Sample 2							
Items	CSR(λ)	USR(λ)	SK(λ)	SJ(λ)	CH(λ)	IS(λ)	MI(λ)	OI(λ)	CSR(λ)	USR(λ)	SK(λ)	SJ(λ)	CH(λ)	IS(λ)	MI(λ)	OI(λ)
SK																
SK5	**.13**		**.55*** **	−.10	.17**	.01	.19**	.02	**.36*** **		**.40*** **	−.08	.15*	.12*	.33***	−.005
SK12	**.23**		**.75*** **	−.09	−.03	.10	.02	.07	**.60*** **		**.52*** **	−.19***	.33***	.23**	.21*	.030
SK19	**.11**		**.68*** **	−.14**	.08	.18**	.13*	.12*	**.45*** **		**.38*** **	−.01	.30***	.13	.23**	.099
SK23	**−.16**		**.52*** **	.64***	.08**	.21*	.13	−.09	**−.02**		**.59*** **	.41**	.07	.19**	.01	.034
SK26	**.09**		**.39*** **	.26***	.23***	.14	.44***	.24***	**.17**		**.40*** **	.29***	.16*	.18**	.29***	.185***
SJ																
SJ1		**−.21**	.04	**.31*** **	.13**	.36***	.05	.11		**−.17**	.18	**.43*** **	−.06	.18**	.13*	.185**
SJ8		**.06**	.14**	**.14***	−.11	.05	−.16**	−.04		**−.14**	−.03	**.24*** **	.06	.34***	−.19**	.074
SJ11		**.06**	.30***	**.47*** **	−.11	.08	.19***	.31***		**.06**	.54***	**.24**	−.14	.22***	.08	.184**
SJ16		**.14**	.15**	**.35*** **	−.07	.29*	−.06	.32***		**.36***	.07	**.34*** **	−.15*	.18	−.05	.500***
CH																
CH3	**−.02**		.18**	−.17*	**.49*** **	.18*	.42**	.13*	**.08**		.22*	−.22***	**.18***	.26***	.24***	−.001
CH7	**.32**		.21***	−.14**	**.46*** **	.02	.06	.07	**−.16**		.29**	−.08	**.54*** **	.04	.16**	.075
CH10	**.63***		.16	.18**	**.55**	−.05	−.23*	.13*	**−.12**		.24	.09	**.73*** **	−.03	.13**	.041
CH15	**.36*** **		.17**	.03	**.35*** **	.18**	.23***	.01	**−.23**		.21*	−.20**	**.38*** **	.08	.31***	−.006
IS																
IS4		**−.03**	.20***	.23***	.06	**.49*** **	.10	.28*		**.02**	.14*	.20	−.07	**.25*** **	.01	.585****
IS13		**.33***	.15**	.09*	.12*	**.43***	.16*	.42***		**.18**	.14*	.04	−.05	**.42*** **	.13*	.413***
IS18		**.21**	.13**	.15***	.01	**.50*** **	.12*	.29***		**.16**	.26***	.17**	.03	**.57*** **	.13**	.159**
IS25		**.26**	.23***	.06	.10	**.58*** **	.21***	.29***		**.07**	.24***	.16***	.15***	**.68*** **	.12**	.353***
MI																
MI9	**.13**		.36***	−.06	.31***	.05	**.30*** **	.25***	**.23***		.40***	−.20	.13	.01	**.36*** **	.385***
MI14	**.35*** **		.39***	.06	.13	.20***	**.41*** **	.13**	**−.09**		.29***	−.08	.11	.20***	**.54*** **	.149**
MI17	**.37***		.24***	.09*	.03	.32***	**.59*** **	.01	**.23**		.10	.184**	.17*	.10	**.76*** **	.040
MI22	**.12**		.38***	−.05	.24***	.15*	**.10**	.01**	**.23**		.19*	−.028	.29***	.07	**.36*** **	.34***
OI																
OI2		**−.49***	.09	.06	.15*	.61***	−.10	**.41**		**−.23**	.04	.128*	.29*	.19***	.12**	**.68*** **
OI6		**.11**	.11*	.26***	.06	.49**	.08	**.42*** **		**.24**	−.05	.28***	.11	.35*	.17**	**.54*** **
OI20		**−.11**	−.02	.26***	.09	.09	.17**	**.60*** **		**−.10**	.16*	.13	−.17**	.24***	.31***	**.40*** **
OI24		**.12**	.21***	−.01	.05	.37*	.09	**.60*** **		**−.18**	.16**	.06	.02	.44***	.16**	**.53*** **

Note. Target factor loadings are in bold. λ = standardized factor loadings. ESEM = exploratory structural equation modeling; SCS = self-compassion scale; CSR = compassionate self-responding; USR = uncompassionate self-responding; SK = self-kindness; SJ = self-judgment; CH = common humanity; IS = isolation; MI = mindfulness; OI = over-identification.
**p* < .05. ***p* < .01. ****p* < .001.

general factor in both samples. Furthermore, for every specific factor, one or two items were not loaded on the hypothesized factor.

The six-factor correlated model (Model 3; Table S6), as mentioned above, was only supported in Sample 2; the six factors were fairly defined, but these factors were not intercorrelated. Self-judgment was not significantly correlated with common humanity and mindfulness, and overidentification was not significantly correlated with common humanity. All 25 items significantly loaded on the general factor, for the single-bifactor model with six specific factors (Model 4; Table S6) in Sample 2, and three specific factors (self-kindness, common humanity, and mindfulness) were identified as hypothesized. However, the negative factors were not defined well. Specifically, there were two items significantly loading onto self-judgment, and there were three items significantly loading onto isolation. However, there was only one item significantly loading onto overidentification.

In summary, although there were more ESEM solutions supported with acceptable model fit in both samples when Item 21 was excluded, the factor loadings as hypothesized were systematically low, and there were still some cross-loadings of some items. The six-factor structure was supported in Sample 2. Although the two-bifactor model with six specific factors (Model 5) has a good model fit, the two general factors, compassionate self-responding and uncompassionate self-responding, were not defined. Thus, the assumption that six specific factors operate in two systems (i.e., compassionate self-responding and uncompassionate self-responding) was not supported. The novel four-factor structure was supported in both samples, especially, the general factor of the bifactor model was defined well in both samples, but the specific factors were not defined consistently in both samples.

The Version of SCS-C Without Item 10: Replicating the Six-Factor Structure in Sample 1. The six-factor correlated model (Model 3) and the single-bifactor model with six specific factors (Model 4) were replicated across languages (Neff et al., 2019), and this model was supported in Sample 2 using ESEM for the version SCS-C without Item 21. However, in Sample 1, the models were not supported for the version of SCS-C without Item 21. As for the full version, in Sample 1, the six-factor correlated model (Model 3) was supported using ESEM with the acceptable model fit, but there was an identification issue in the single-bifactor model with six specific factors (Model 4). We checked the Mplus warning message and found that the problematic item in Sample 1 for identifying the single-bifactor model with six specific factors (Model 4) was Item 10 ("When I feel inadequate in some way, I try to remind myself that feelings of inadequacy are shared by most people") from common humanity. In order to replicate the single-bifactor model with six specific factors (Model 4) in Sample 1, we ran the analysis of the six-factor correlated model (Model 3) and the single-bifactor model with six specific factors (Model 4) using the SCS-C only without Item 10 (Table S7).

After removing Item 10, the identification issue was solved for the single-bifactor model with six specific factors (Model 4) both using CFA and ESEM. For the six-factor correlated model (Model 3) using CFA, it was not supported with acceptable model fit (χ^2 = 1,157.87, df = 26, $p < .001$, CFI = .85, TLI =.82, RMSEA [90% C.I.] = .09 [.08, .09]), SRMR = .062, but the six-factor correlated model (Model 3) using ESEM was supported by the acceptable model fit (χ^2 =309.44, df = 165, $p < .001$, CFI = .98, TLI =.96, RMSEA [90% C.I.] = .04 [.04, .05]), SRMR = .02. We found cross-loading problems for some items and identified three items that did not significantly load on the hypothesized factors: for self-judgment, Items 1 ("I'm disapproving and judgmental about my own flaws and inadequacies") and 11 ("I'm intolerant and impatient toward those aspects of my personality I don't like"), and for mindfulness, Items 9 ("When something upsets me I try to keep my emotions in balance") and 22 ("When I'm feeling down I try to approach my feelings with curiosity and openness"). As for the correlation between factors, self-judgment was not significantly correlated with self-kindness (r = .12) and mindfulness (r = .004), and common humanity was not significantly correlated with isolation (r = .07), mindfulness (r = .011) and overidentification (r = −.02).

The single-bifactor model with six specific factors (Model 4; Table S7) using CFA was not supported with the poor model fit (χ^2 = 1,038.38, df = 250, $p < .001$, CFI = .87, TLI =.84, RMSEA [90% CI] = .08 [.08, .09]), SRMR = .07, whereas the single-bifactor model with six specific factors (Model 4) using ESEM was supported with the acceptable model fit (χ^2 = 252.72, df = 146, $p < .001$, CFI = .98, TLI =.96, RMSEA [90% CI] = .04 [.03, .05], SRMR = .02). The general factor of self-compassion was fairly defined (λ = .26–.72, $M\lambda$ = .50), but only 23 items loaded onto the general factor, and two items from self-judgment, Items 8 ("When times are really difficult, I tend to be tough on myself") and 21 ("I can be a bit cold-hearted toward myself when I'm experiencing suffering") did not significantly load onto the general factor. As for the specific factors, Items 23 ("I'm tolerant of my own flaws and inadequacies") and 26 ("I try to be understanding and patient toward those aspects of my personality I don't like") did not load onto self-kindness as hypothesized. For self-judgment, Items 1 and 11 did not load; for common humanity, all three Items 3, 7, and 15 loaded, λ = .41/.44/.38. Four items significantly loaded onto isolation as hypothesized (λ = .46–.43, $M\lambda$ = .352). However, all four items did not significantly load onto mindfulness. Item 20 ("When something upsets me, I get carried away with my feelings") did not significantly load on overidentification. We also found there are cross-loadings for some items.

In summary, based on the six-factor correlated model (Model 3), we found the factors, self-judgment and mindfulness in Sample 1 were problematic. Based on the findings related to the single-bifactor model with six specific factors (Model 4), for the general factor, two items from self-judgment did not significantly load on the general factor. As for the specific factors, self-kindness, self-judgment, and mindfulness seemed to be problematic.

Bifactor Model Evaluation

In total, two single-bifactor models were supported in different versions of SCS-C, the single-bifactor model with four specific factors using SCS-C omitting Item 21 in both samples, and the single-bifactor model with six specific factors using SCS-C omitting Item 10 in Sample 1 and Item 21 in Sample 2. We evaluated bifactor models which were supported in both samples based on Rodriguez et al. (2016b).

For the single-bifactor model with four specific factors (Model 7), omega for the general factor was high, omega = .91/.91, Sample 1/2, but omegaH was low in both samples, omegaH =.73/.66, Sample 1/2. The ratio of the omegaH/omega was 79.85% in Sample 1 and 72.43% in Sample 2, indicating that around 80%/72% (Sample 1/2) of reliable variance in the total score can be attributed to the general factor self-compassion (Rodriguez et al., 2016b). In contrast, only around 18%/25% (Sample 1/2) of the reliable variance ($\omega - \omega_H$) in the total score can be attributed to the multidimensionality caused by the four specific factors, and 9% ($1-\omega$) in both samples could be estimated to be due to random error. The construct reliability (H) index for the general factor (self-compassion) was .89/.90, Sample 1/2, indicating that the general factor was represented well by the 26 items and that the general factor suggested a well-defined latent variable across both samples. The FD for the general factor was .94/.96 in Sample 1/2, indicating that the factor score from the general factor (self-compassion) is trustworthy, and it allows the factor score estimation of this general factor in SEM research.

For the single-bifactor model with six specific factors (Model 4), in Sample 1, we replicated this model using the version of SCS-C without Item 10. The omega was .92, and omegaH was .84. Thus, around 91% of reliable variance in total score can be attributed to the general factor, self-compassion. Only around 8% of the reliable variance in the total score can be attributed to the multidimensionality caused by the six specific factors, and 8% could be estimated to be due to random error. Both H (.90) and FD (.95) were high. These findings indicated that the general factor, but not individual subscales, was represented well by the 25 items, and it can be used in later SEM context. In Sample 2, we replicated the model using the version of SCS-C without Item 21. Evaluation of the general factor revealed omega = .92, omegaH =. 80, H = .90, and FD =.96, were all above the benchmarks indicating that the general factor, but not individual subscales, was represented well by the 25 items, and it can be used in the SEM context.

Discussion

Our research aimed to address problems and gaps around the factor structure of the Chinese version of the SCS. Informed by previous research, we examined seven different models and applied recent statistical analysis recommendations (i.e., ESEM) to investigate whether previous failures of establishing the original factor structure in Chinese samples could be accounted for by approaches that do not consider the hypothesized synergetic nature of six components of self-compassion. We also explored whether the SCS-C supports the understanding of self-compassion as one factor or two factors, compassionate self-responding and uncompassionate self-responding (e.g., Brenner et al., 2017; Costa et al., 2016; López et al., 2015; Neff et al., 2019). We first investigated the full version of SCS-C in two samples of Chinese young adults and then repeated our analysis omitting particularly problematic items. Finally, we evaluated the variance explained by the general factor for those bifactor models supported by acceptable model fits (i.e., single-bifactor model with four or six specific factors).

For the full SCS-C, none of our analyses unequivocally supported the originally proposed six-factor structure (Models 3–5; Neff, 2003) in both samples. Although the six-factor structure of the SCS-C has been replicated in some Chinese studies others failed to replicate it (Tsai, 2015; Zeng et al., 2016). These findings, taken together with our findings, suggest that in Chinese samples, the factor structure of the SCS might not be in line with the original six-factor structure definition. This is surprising, given that the six-factor correlated model has been replicated in many samples and languages (Neff et al., 2019).

The possibility that the SCS in Chinese samples may follow a novel four-factor structure as proposed in Tsai (2015) was also only partially supported with acceptable fits using ESEM in both samples; all four factors were defined with at least medium size factor loadings in only Sample 2, whereas in Sample 1, mindfulness was not specified at all. In addition, for its bifactor model with one general factor, we only obtained acceptable model fits in Sample 1, whereas we had a model identification issue in Sample 2. The potential explanation for this issue in the sample may be due to the low intercorrelation between the four factors. Thus, although the four-factor correlated model was supported in both samples, this solution is not a stable construct across both samples due to these factor specification issues.

We further explored the factor structure of SCS-C without problematic Item 21 ("I can be a bit cold-hearted toward myself when I'm experiencing suffering") from self-judgment and reran the whole analysis. The six-factor correlated model (Model 3) was not supported in Sample 1 anymore but supported with good model fits in Sample 2, and all six factors were specified. Also, the single-bifactor model with six specific factors (Model 4) was supported in Sample 2 with acceptable model fits. The novel four-factor structure (Models 6 and 7) was supported in both samples with acceptable model fits, despite some issues of factor specification found in the full version of SCS-C.

Taken together, our findings suggest that the SCS-C may have a different factor structure in Chinese samples than proposed by the initial theory (Neff, 2003, 2016) Still, evidence for an alternative four-factor structure was also not unequivocal.

After failing to fully resolve the psychometric problems with SCS-C, we proposed that Chinese participants may have a different understanding of the components of SCS-C. It may be useful to explore this in more detail to unpack some unexpected findings of our study further. Based on the recent focus group study (Zhao, Smithson et al., 2021), particular discrepancies were identified in the perception of common humanity, which was partially described as a "negative coping strategy" that resembles an excuse for failure. In contrast, the perceptions of self-judgment and over-identification were partially described as helpful, with "self-criticism as an adaptive strategy" and "helpful for self-reflection." These alternative views may guide participants to answer the questions from a different perspective to the western samples on which the scale was initially developed.

The problems observed with the mindfulness subscale could be attributed to challenges in translating the term so that it resonates with the core concept of mindfulness. When translating the self-compassion scale, we found that in Chinese, there is no literal translation for the combination of the words "emotion" and "balance" to express "keep emotion in balance" (保持情绪平衡). Instead, the contemporary Chinese language uses "keep emotion stable" (保持情绪稳定), which indicates both emotional stability and balance while being aware of the emotion. Native English speakers who are not experts in clinical psychology confirmed that "keep emotion stable" is a synonym for "keep emotion in balance." Therefore, we chose the translation of "keep emotion stable" while being aware of the potential discrepancy between the two languages. Our study highlighted that sometimes it is hard to find an equivalent recognized term in a different language that adequately captures the underlying concept. For example, a participant mentioned that there was no translation for the word "optimistic" in Pashtun in a focus group study aiming to discuss the items from the Warwick-Edinburgh mental well-being scale (Taggart et al., 2013). Thus, it is important to be aware of such language discrepancies when translating scales and evaluating a translated version scale.

Although we found partial support for both the four-factor structure and the six-factor structure with fairly acceptable model fits and factor loadings, given the issues with factor specifications, our psychometric results do not support one specific factor structure over another. The finding from a qualitative focus group study about the understanding of self-compassion (Zhao, Smithson et al., 2021) regarding individuals' views of the negative subscales was in line with the psychometrically suggested general negative component. However, further research on different samples is necessary to confirm support for this. If a four-factor model would be endorsed, it is in contrast to the psychometrically better established original six-factor structure (Neff et al., 2019). Our and Neff et al.'s (2019) recent studies highlighted the need for more research to inform a revised SCS-C.

Furthermore, for the bifactor model of six specific factors, in contrast to Neff et al. (2019), our results suggest that the use of the SCS-C total score is not strongly supported. In their study, both high omega and omegaH were identified. We only replicated the high omega but not high omegaH, which reflects the proportion of variance in total scores attributable to a single general factor (Reise, Moore et al., 2013). Therefore, we need to be cautious to suggest using the total score in our SEM because only when both omega and omegaH are high, as is the case in Neff et al.'s (2019) study, the total score can be calculated. The rationale for this recommendation is that if the value of omegaH is not as high as omega, the difference of the total score could be driven by a specific factor rather than the general factor, which could lead to a misinterpretation of the general factor (Rodriguez et al., 2016a). Taken together, based on our findings, we would, therefore, not suggest using the total score of the general factor of SCS-C without a relevant measurement check.

Similarly, the two-bifactor model with six specific factors (Model 5) was supported, but the items predominantly did not load significantly onto both of the general factors. Some researchers provided evidence that self-compassion could be operationalized as two separate total scores, compassionate self-responding and uncompassionate self-responding, rather than one total score (Brenner et al., 2017; Costa et al., 2016; Halamová et al., 2020; López et al., 2015). Our findings did not support this, and our previous work shows that when reflecting on the six specific factors of self-compassion, Chinese participants reported both positive and negative perspectives of each factor, suggesting that they perhaps adopt a dialectical approach in their understanding of the factors (Zhao, Smithson et al., 2021). This could explain why in our Chinese samples, we did not find strong empirical support for separating positive and negative factors. However, limited empirical studies in Chinese samples have explored this debate. We, therefore, suggested future studies focus on exploring this issue.

The scoring of the self-compassion scale, such as scoring as a total self-compassion score (e.g., Neff et al., 2019) or as two separate scores (Brenner et al., 2017; Costa et al., 2016; Halamová et al., 2020; López et al., 2015), and the factor structure of SCS remain unresolved. Our findings suggested that it is hard to replicate models across samples and that the Chinese version of the SCS, despite careful language considerations, does not work as well as it works for translations into other languages. However, there has been increasing interest in self-compassion, and SCS-C is currently the only measure for assessing self-compassion in China.

In the interim, before a valid SCS is available for Chinese participants, researchers should consider two points when using the current SCS-C. First, although there is limited evidence supporting the single general factor with acceptable model fits, we do not encourage researchers to sum all items together to get a total score. Instead, we strongly recommend that researchers use a latent variable approach for their measurement construct in the SEM context as the findings of FD and construct reliability demonstrated. Second, we caution against using individual subscale scores of the SCS-C given the unclear factor structure, systematically low factor loadings, and problematic factor specification. We are aware that this is not convenient for future studies in Chinese samples with smaller sample sizes, but based on our recent findings, together with those from Neff et al. (2019), there is limited evidence for the use of the total score of SCS-C or individual subscales in research.

Theoretical Implications and Future Direction

Our findings suggested that it is necessary to revise the SCS-C in order to account for the possibility of a different factor structure. Groups of participants should be invited to freely discuss the translation and meaning of the items to make sure that the items assess the same scenario and avoid language discrepancy, such as the discrepancy of "keep emotion in balance" (Zhao, Smithson et al., 2021), which may help to develop a clear six-factor structure. Furthermore, larger-scale qualitative research (Smithson, 2020) could help confirm the themes and might inform whether the SCS-C in its current form needs additional items to help differentiate the negative factors in this population. For example, for self-judgment, qualitative evidence supported "self-criticism as undermining the self" and "self-criticism as an adaptive strategy" of which the existing items cover the former but not the latter description. Thus, extra items that describe benign self-criticism may potentially be added. For example, "I try to reflect on the failure to overcome my flaws when I fail in something important." Future studies should continue to explore the understanding and meaning of the main components of self-compassion in Chinese populations as well as the understanding of specific items.

Further research on the construct in Chinese populations could draw on other aspects proposed to characterize compassion, such as engaging in one's suffering with an open mind and taking wise actions to handle it (Gilbert et al., 2017). In a recent review, five components of compassion were identified: recognition of suffering; understanding its universality; feeling sympathy, empathy, or concern for those who are suffering; tolerating the distress associated with the witnessing of suffering; and motivation to act or acting to alleviate the suffering (Strauss et al., 2016). The action of handling current difficulties appears to be a core component that could be added to the questionnaire to assess self-compassion in Chinese samples. We do not argue that involving the action component is more important than the other components, but there is a published SCS based on Strauss et al.'s (2016) definition (Gu et al., 2020), and it was revealed in a preliminary focus group study that in Chinese samples self-compassion appears to be associated with active problem-solving in difficult times, for example, "problem-solving" and "self-reflection" (Zhao, Smithson et al., 2021).

Finally and ideally, in connection with the previous suggestions, we may need to extend the self-compassion definition by embedding theories relevant to the Chinese culture as specific features of self-compassion may vary by cultural background (Montero-Marin et al., 2018; Neff et al., 2008; Zhao, Smithson et al., 2021). In addition, the SCS was not supported in Japanese and Chinese samples, which both are collectivist cultures (Oyserman et al., 2002). As defined in the self-construal theory by Markus and Kitayama (1991), collectivist cultures are characterized by a dominant interdependent self-construal and encouraging social harmony, thus shaping individuals' emotion, motivation, thinking style and behaviors and also their understanding of self-compassion. For example, individuals with high interdependent self-construal tend to use self-criticism as a motivation to grow to achieve social harmony, thus, Chinese participants may not consider self-judgment as an uncompassionate behavior. Second, Confucianism could be another lens through which self-compassion can be considered in China. There are many critical values in Confucianism still dominant in the current society. For example, Confucianism encourages introspection or self-reflection as the main way for personal growth and for improving the relationship between self and others (Cheng, 2004). Furthermore, in Confucianism, shame is considered a positive motivation (Mencius, 372BC-289BC; Seok, 2015), a perspective different from Western psychology, where shame was considered as a source of negative self-evaluation and a maintenance factor for mental health problems (Gilbert, 2003). This specific Confucianism mentality could also be considered as a form of interdependent self-construal. Taken together, drawing self-compassion into the specific cultural context is not only beneficial for its assessment but also promoting relevant cultural adaptation of psychological therapy.

Limitations

This study was designed to conduct a comprehensive investigation of the psychometric properties of SCS-C using both CFA and ESEM, but there are several limitations. First, this study sample was university students, and the findings may not generate in other groups, such as clinical patients. In addition, our samples are dominated by women, so the findings may not represent for men or gender non-conforming individuals. Another issue is that when requesting demographic information, we used "xingbie" in Chinese. Xingbie is a generic word that can mean gender or sex. Thus, we used participants' self-report of sex/gender and cannot be sure whether they are referring to sex assigned at birth or gender. We suggest future studies can use a more specific word to collect data to avoid the issue. Third, this report does not include other validity tests, such as convergent and discriminate validity. Although we acquired this information from all participants similarly to Neff's original validation study (2003), we abandoned its further evaluation because we did not get consistent findings related to the factor structure. Furthermore, the two samples in our study were recruited via two different formats (online vs paper and pen). It is possible that assessment context (e.g., paper-and-pen format was assessed during a university lecture vs online was completed at participants' choice in their own time) introduced between-sample differences. We would suggest future studies using the same way to recruit participants to reduce the potential between-sample differences. Last, we have model misspecification issues in some models across samples. Although the sample size is associated with model misspecification, the sample sizes in the current study are comparable to or somewhat larger than previous SCS validation studies. The bifactor-ESEM model is complex, thus, our study could lack sufficient statistical power. We suggest future SCS-C studies using a larger sample size to replicate the relevant models.

Conclusion

We suggest that the SCS-C needs to be revised, and there may be a cultural difference in the understanding of self-compassion in the Chinese sample. When using the existing SCS-C in research, especially for path models, researchers are advised to use a latent variable approach and establish the measurement construct first rather than sum scores of the full scale or subscales without checking psychometric properties.

Acknowledgments

This work was supported by the Research & Training Support Grant (RTSG) funds, College of Life and Environmental Sciences (CLES), the University of Exeter. The authors gratefully acknowledge Professor Jianing You at the South China Normal University for providing the scale. We thank all research assistants for helping recruitment, especially Yuanlin Li. We gratefully acknowledge all researchers and clinical practitioners for helping with the translation, Dr Kai He, Yao Lu at the Beijing Foreign Study University, Alicia Cork at the University of Exeter, Jiexuan Li, Mingze Wang at the University of Auckland. The authors gratefully acknowledge the discussion regarding statistics Tuo Liu at the University of Oldenburg and Dr Yang Liu at the University of Maryland. The authors gratefully acknowledge Dr Lindsay Walker and Dr Channing Mathews for proofreading the manuscript.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Ethics Approval

The study was approved by the Psychology Ethics Committee of the University of Exeter (United Kingdom, reference number, eCLESPsy000310 v2.1).

ORCID iD

Mengya Zhao https://orcid.org/0000-0001-8078-6514

Availability of Data and Analysis Syntax

All data and data analysis syntax are available upon request from the authors.

References

*References marked with an asterisk were used in the meta-analysis.

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 397–438.

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, *16*(1), 74–94.

Brenner, R. E., Heath, P. J., Vogel, D. L., & Credé, M. (2017). Two is more valid than one: Examining the factor structure of the Self-Compassion Scale (SCS). *Journal of Counseling Psychology*, *64*(6), 696–707.

Brenner, R. E., Vogel, D. L., Lannin, D. G., Engel, K. E., Seidman, A. J., & Heath, P. J. (2018). Do self-compassion and self-coldness distinctly relate to distress and well-being? A theoretical model of self-relating. *Journal of Counseling Psychology*, *65*(3), 346–357.

Chen, J., Yan, L-s, & Zhou, L.-h. (2011). Reliability and validity of Chinese version of self-compassion scale. *Chinese Journal of Clinical Psychology*, *19*(6), 734–736.

*Cheng, R. H. (2004). Moral education in Hong Kong: Confucian-parental, Christian-religious and liberal-civic influences. *Journal of Moral Education*, *33*(4), 533–551.

Costa, J., Marôco, J., Pinto-Gouveia, J., Ferreira, C., & Castilho, P. (2016). Validation of the psychometric properties of the self compassion scale. Testing the factorial validity and factorial invariance of the measure among borderline personality disorder, anxiety disorder, eating disorder and general populations. *Clinical Psychology & Psychotherapy*, *23*(5), 460–468.

Crocker, J., & Canevello, A. (2008). Creating and undermining social support in communal relationships: The role of compassionate and self-image goals. *Journal of Personality and Social Psychology*, *95*(3), 555–575.

Deng, P., Ding, L., Zhang, Y., & Deng, Y. (2012). Reliability and validity of revised version of self-compassion scale in youth (修订版自我怜悯量表在青少年中应用的信效度). *Journal of Campus Life & Mental Health*, *10*(6), 371–373.

Dueber, D. (2020). *Bifactor indices calculator: Bifactor indices calculator* (Version 0.2.0). R Package.

Germer, C. K., & Neff, K. D. (2013). Self-compassion in clinical practice. *Journal of Clinical Psychology*, *69*(8), 856–867.

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.

Gilbert, P. (2003). Evolution, social roles, and the differences in shame and guilt. *Social Research: An International Quarterly*, *70*(4), 1205–1230.

Gilbert, P. (2005). *Compassion: Conceptualisations, research and use in psychotherapy*. Routledge.

Gilbert, P., Catarino, F., Duarte, C., Matos, M., Kolts, R., Stubbs, J., . . . Basran, J. (2017). The development of compassionate engagement and action scales for self and others. *Journal of Compassionate Health Care*, *4*(1), Article 4.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Erlbaum.

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.

Halamová, J., Kanovský, M., Petrocchi, N., Moreira, H., López, A., Barnett, M. D., . . . Zengar, M. (2020). Factor structure of the self-compassion scale in 11 international samples. *Measurement and Evaluation in Counseling and Development*, *54*, 1–23.

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. du Toit, & D. Sorbom (Eds.), *Structural equation modeling: Present and future* (pp. 195–216). Scientific Software International.

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jiang, Y., You, J., Ren, Y., Sun, R., Liao, S., Zhu, J., & Ma, N. (2017). Brief report: A preliminary comparison of self-compassion between adolescents with nonsuicidal self-injury thoughts and actions. *Journal of Adolescence*, *59*, 124–128.

Jiang, Y., You, J., Zheng, X., & Lin, M.-P. (2017). The qualities of attachment with significant others and self-compassion protect adolescents from non suicidal self-injury. *School Psychology Quarterly*, *32*(2), 143–155.

Jing, K., Wang, J., & Liu, F. (2011). Reliability and validity of self-compassion scale (大学生自我怜悯问卷的修订及信效度研究). *Socialpsychological Sicence (*社会心理科学*)*, *26*(8), 937–940.

Körner, A., Coroiu, A., Copeland, L., Gomez-Garibello, C., Albani, C., Zenger, M., & Brähler, E. (2015). The role of self-compassion in buffering symptoms of depression in the general population. *PLOS ONE*, *10*(10), Article e0136598.

López, A., Sanderman, R., Smink, A., Zhang, Y., Van Sonderen, E., Ranchor, A., & Schroevers, M. J. (2015). A reconsideration of the self-compassion scale's total score: Self-compassion versus self-criticism. *PLOS ONE*, *10*(7), Article e0132940.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*(2), 224–253.

Marsh, H. W., Hau, K.-T., & Grayson, D. (2005). Goodness of fit in structural equation models. In A. Maydeu-Olivares, & J. J. McArdle (Eds.), *Contemporary psychometrics: A festschrift for Roderick P. McDonald* (pp. 275–340). Erlbaum.

Marsh, I. C., Chan, S. W., & MacBeth, A. (2018). Self-compassion and psychological distress in adolescents—A meta-analysis. *Mindfulness*, *9*(4), 1011–1027.

Montero-Marin, J., Kuyken, W., Crane, C., Gu, J., Baer, R., Al-Awamleh, A. A., . . . García-Campayo, J. (2018). Self-compassion and cultural values: A cross-cultural study of self-compassion using a multitrait-multimethod (MTMM) analytical procedure. *Frontiers in Psychology*, *9*, Article 2638.

Morin, A. J. S., Marsh, H. W., & Nagengast, B. (2013). Exploratory structural equation modeling. In G. R. Hancock, & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 395–436). Information Age.

Muris, P., & Otgaar, H. (2020). The process of science: A critical evaluation of more than 15 years of research on self-compassion with the self-compassion scale. *Mindfulness*, *11*(6), 1469–1482.

Neff, K. D. (2003). The development and validation of a scale to measure self-compassion. *Self and Identity*, *2*(3), 223–250.

Neff, K. D. (2016). The self-compassion scale is a valid and theoretically coherent measure of self-compassion. *Mindfulness*, *7*(1), 264–274.

Neff, K. D., Pisitsungkagarn, K., & Hsieh, Y.-P. (2008). Self-compassion and self-construal in the United States, Thailand, and Taiwan. *Journal of Cross-cultural Psychology*, *39*(3), 267–285.

Neff, K. D., Tóth-Király, I., Yarnell, L. M., Arimitsu, K., Castilho, P., Ghorbani, N., . . . Mantzios, M. (2019). Examining the factor structure of the self-compassion scale in 20 diverse samples: Support for use of a total score and six subscale scores. *Psychological Assessment*, *31*(1), 27–45.

Neff, K. D., Whittaker, T. A., & Karl, A. (2017). Examining the factor structure of the self-compassion scale in four distinct populations: Is the use of a total scale score justified? *Journal of Personality Assessment*, *99*(6), 596–607.

Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, *128*(1), 3–72.

Piedmont, R. L. (2014). Inter-item correlations. In A.C. Michalos (Ed.), *Encyclopedia of quality of life and well-being research* (pp. 3303–3304). Springer.

Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, *95*(2), 129–140.

Reise, S. P., Moore, T. M., & Haviland, M. G. (2013). Applying unidimensional item response theory models to psychological data. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J.-I. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbook of testing and assessment in psychology: Vol. 1. Test theory and testing and assessment in industrial and organizational psychology* (pp. 101–119). American Psychological Association.

Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika*, *74*(1), Article 145.

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016a). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of personality assessment, 98*(3), 223–237.

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016b). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137–150.

Seok, B. (2015). Moral psychology of shame in early Confucian philosophy. *Frontiers of Philosophy in China*, *10*(1), 21–57.

Smithson, J. (2020). *Group interviews*. SAGE.

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.

Taggart, F., Friede, T., Weich, S., Clarke, A., Johnson, M., & Stewart-Brown, S. (2013). Cross cultural evaluation of the Warwick-Edinburgh mental well-being scale (WEMWBS)-a mixed methods study. *Health and Quality of Life Outcomes*, *11*(1), Article 27.

Tsai, M.-Y. (2015). Construction and factorial validation of the Chinese version of the self-compassion scale for gifted students. *Psychology*, *5*(11), 634–644.

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.

Yang, X. (2016). Self-compassion, relationship harmony, versus self-enhancement: Different ways of relating to well-being in Hong Kong Chinese. *Personality and Individual Differences*, *89*, 24–27.

Yang, Y., Zhang, M., & Kou, Y. (2016). Self-compassion and life satisfaction: The mediating role of hope. *Personality and Individual Differences*, *98*, 91–95.

Yarnell, L. M., & Neff, K. D. (2013). Self-compassion, interpersonal conflict resolutions, and well-being. *Self and Identity*, *12*(2), 146–159.

Zeng, X., Wei, J., Oei, T. P., & Liu, X. (2016). The self-compassion scale is not validated in a Buddhist sample. *Journal of Religion and Health*, *55*(6), 1996–2009.

Zessin, U., Dickhäuser, O., & Garbade, S. (2015). The relationship between self-compassion and well-being: A meta analysis. *Applied Psychology: Health and Wellbeing*, *7*(3), 340–364.

Zhao, M., Ford, T., Wang, P., & Karl, A. (2021). The role of parenting, self-compassion and friendships in depressive symptoms among young people in the UK and China. *Current Psychology*.

Zhao, M., Smithson, J., Ford, T., Wang, P., Wong, N. Y. B., & Karl, A. (2021). Self-compassion in Chinese young adults: Specific features of the construct from a cultural perspective. *Mindfulness*, *12*, 2718–2728.