AMERICAN PSYCHOLOGICAL ASSOCIATION

Psychological Assessment

ISSN: 1040-3590

2023, Vol. 35, No. 8, 674–691
https://doi.org/10.1037/pas0001247

The Bifactor Structure of the Self-Compassion Scale: Bayesian Approaches to Overcome Exploratory Structural Equation Modeling (ESEM) Limitations

Herbert W. Marsh[1], Madeleine I. Fraser[2], Arman Rakhimov[3], Joseph Ciarrochi[1], and Jiesi Guo[1]
[1] Institute for Positive Psychology and Education, Australian Catholic University
[2] School of Behavioural Health Sciences, Australian Catholic University
[3] International School of Economics, M. Narikbayev KAZGUU University

The rapidly expanding self-compassion research is driven mainly by Neff's (2003a, 2003b, 2023) six-factor Self-Compassion Scale (SCS). Despite broad agreement on its six-first-order factor structure, there is much debate on SCS's global structure (one- vs. two-global factors). Neff et al. (2019) argue for an exploratory structural equation model (ESEM) with six specific and one global bifactor (6ESEM + 1GlbBF) rather than two global factors (6ESEM + 2GlbBF). However, ESEM's methodological limitations precluded testing the appropriate 6ESEM + 2GlbBF, relying instead on a model combining ESEM and traditional confirmatory factor analysis (6ESEM + 2CFA). Although intuitively reasonable, this alternative model results in internally inconsistent, illogical interpretations. Instead, we apply recent advances in Bayesian SEM frameworks and Bayes structural equation models fit indices to test a more appropriate bifactor model with two global factors. This model (as does 6CFA + 2GlbBF) fits the data well, and correlations between compassionate self-responding (CS) and reverse-scored uncompassionate self-responding (RUS) factors (~.6) are much less than the 1.0 correlation implied by a single bipolar factor. We discuss the critical implications for theory, scoring, and clinical application for the SCS that previously were inappropriately based on this now-discredited 6ESEM + 2GlbCFA. In applied practice, we endorse using scores representing the six SCS factors, total SCS, and CS and RUS components rather than relying solely on one global factor. Our approach to these issues (dimensionality, factor structure, first-order and higher order models, positive vs. negatively oriented constructs, item-wording effects, and alternative estimation procedures) has wide applicability to clinical measurement (see our annotated bibliography of 20 instruments that might benefit from our approach).

Public Significance Statement

Neff's (2003a, 2003b, 2023) six-factor Self-Compassion Scale (SCS) dominates self-compassion research. However, there is much debate on SCS's global structure (one- vs. two-global factors), and the scores used to summarize SCS responses. Our critique of ongoing SCS research shows that inappropriate statistical models have resulted in erroneous conclusions at the heart of these debates. Issues we address in our critique have broad generalizability to psychological assessment studies.

Keywords: self-compassion, exploratory structural equation models (ESEM), bifactor structural equation models, Bayes structural equation models (BSEM), substantive-methodological synergy

Supplemental materials: https://doi.org/10.1037/pas0001247.supp

Dimensionality and factor structure are important issues for developing and applying most clinician instruments. For clinical instruments, it is typical to have one or more overarching theoretical constructs AND multiple specific factors. For example, Wechsler's (2014) widely used intelligence test posits one overarching global construct (IQ) and specific constructs (e.g., verbal comprehension, visual–spatial, fluid reasoning, working memory, and processing speed). Likewise, Larsen and Diener's (1987) Affect Intensity

This article was published Online First July 6, 2023.
Herbert W. Marsh https://orcid.org/0000-0002-1078-9717
Madeleine I. Fraser https://orcid.org/0000-0002-1824-6661
Arman Rakhimov https://orcid.org/0000-0001-7501-4317
Joseph Ciarrochi https://orcid.org/0000-0003-0471-8100
Jiesi Guo https://orcid.org/0000-0003-2102-803X
The authors would like to acknowledge the support for Kirstin Neff and István Tóth-Király for providing secondary anonymous data used in the present investigation, to Kirstin Neff for helpful comments at all stages of the research, to Alex Morin for input into the statistical models used in the present investigation, and to Bengt Muthén and Tihomir Asparouhov for advice on the application of Mplus statistical package—particularly limitations on the use of exploratory structural equation modeling for bifactor models with two global factors.

This study was based on a reanalysis of data from Neff et al. (2019) that were graciously provided by the Kristin Neff. Enquiries about further

continued

Measure (AIM) has two overarching global higher order factors (positive affectivity and negative affectivity) and six first-order specific factors (fear, sadness, anger, guilt, joy, and love). Depending on the instrument's intent and the application, the main focus might be on the specific factors or the overarching construct. However, in many instances, both global and specific factors are relevant. Nevertheless, this balance between global and specific factors is often a source of contentious debate. For example, if the global factor is sufficiently strong, the specific factors might be too highly correlated to be meaningfully differentiated. Getting the balance right is important for many clinical instruments (see supplemental material Section 1).

In addition to positing both global and specific factors, clinical instruments often contain a mixture of positively oriented and negatively oriented (negatively worded or reverse-scored) items. Larsen and Diener's (1987) AIM instrument mentioned earlier is one example. However, positively and negatively oriented items almost always load on different factors (or provide marginal fits when forced to load on a single factor). Higher order or bifactor models positing two global factors are viable approaches to this issue. These can be posited based on substantive theory (i.e., meaningfully different constructs) or nonsubstantive method effects (e.g., item-wording effects). For example, the Basic Needs Satisfaction and Frustration Scale (Chen et al., 2015) assesses whether basic psychological needs for autonomy, competence, and relatedness are satisfied (positively oriented) or frustrated (negatively oriented). The six specific factors are combined to form two global factors (need satisfaction and need frustration). Here the global positively and negatively oriented global factors are substantively meaningful factors based on theory. However, sometimes it is difficult to determine whether separate positively and negatively oriented factors are substantively important or an item-wording artifact. How best to deal with positive and reverse-scored items is an essential issue for many clinical instruments (see supplemental material Section 1).

Competing factor analysis models used to test dimensionality include: unidimensional models (ignoring the multiple specific factors), correlated trait models (ignoring the global factor), higher or second order models (with global higher order and multiple specific factors), and bifactor models (with global bifactor and multiple specific factors). Higher order and bifactor models both contain a mixture of specific and global factors. In higher order models, the higher order (HO) global factors are defined in relation to the first-order factors. In bifactor models, the global bifactors (GlbBF) are defined by the individual items (see subsequent discussion).

Clinical researchers typically test factor structures with confirmatory factor analysis (CFA). CFA models usually start with an a priori structure in which every item loads on one and only one factor. However, this assumption is often too restrictive in applied clinical research to achieve adequate goodness of fit. Instead, items tend to have minor cross-loadings on other factors, even if they load mainly on the factor they are designed to measure. Also, when small nonzero cross-loadings are constrained to zero, the relations between factors are likely positively biased. Thus, CFA models are highly parsimonious but typically overly restrictive for most clinical instruments.

Exploratory structural equation modeling (ESEM; Marsh et al., 2014) with target rotation is increasingly used in clinical research. This alternative combines many of the best features for CFA and exploratory factor analysis (EFA). Like CFA, it starts with an a priori model. However, like EFA, it allows items to cross-load on different factors. Comparing the fit and parameter estimates based on ESEM and CFA is informative. If the fit and parameter estimates are similar, then CFA is preferred on the basis of parsimony. However, if ESEM fits the data better and the parameter estimates are more in line with the a priori theoretical structure, ESEM is preferred—as will often be the case.

Bayes estimation represents an alternative to ESEM. It is infrequently used in applied clinical research due in part to its technical difficulty. However, many of these issues have been overcome as more user-friendly applications evolve, making it easier to use and providing goodness of fit indices like those available in CFA and ESEM. Bayes is particularly useful for a few models that cannot readily be implemented with ESEM.

These issues (dimensionality, factor structure, first-order and higher order models, positive vs. negatively oriented constructs, item-wording effects, and alternative estimation procedures) have wide applicability to clinical measurement. Thus, many clinical and psychological assessment instruments have had to deal with these concerns (see annotated bibliography in supplemental material Section 1). Here we explore these issues in a critique of recent and ongoing research into the construct of self-compassion as measured by Neff's (2003a, 2003b, 2023) Self-Compassion Scale (SCS). Our study is a substantive-methodological synergy, integrating the application of strong, new, or evolving methodology that more appropriately addresses clinically substantive issues that have implications for theory, policy, and practice. Here we address the appropriateness of statistical methods in the widely debated issue of whether overall self-compassion represents a single (bipolar) factor or two (positive and negative) factors. Although we focus on self-compassion, the issues have broad generalizability to clinical measurement.

access to the data should be sent to Kristin Neff (kristin@self-compassion.org). In the present investigation, code for analyses is provided in Neff et al.'s (2019) study and in supplemental materials. The present investigation was not preregistered. Because the study was a reanalysis of secondary anonymous data previously published in Psychological Assessment (Neff et al., 2019) and generously provided by the lead author of that article, it was deemed not to require further ethical approval.

Herbert W. Marsh played a lead role in conceptualization, formal analysis, investigation, methodology, writing–original draft, and writing–review and editing. Madeleine I. Fraser played a supporting role in conceptualization, investigation, methodology, writing–original draft, and writing–review and editing. Arman Rakhimov played a lead role in visualization and a supporting role in conceptualization, formal analysis, methodology, and writing–review and editing. Joseph Ciarrochi played a supporting role in conceptualization, methodology, writing–original draft, and writing–review and editing. Jiesi Guo played a supporting role in conceptualization, methodology, and writing–review and editing.

Correspondence concerning this article should be addressed to Herbert W. Marsh, Institute for Positive Psychology and Education, Australian Catholic University, 33 Berry Street, North Sydney, NSW 2060, Australia. Email: Herb.marsh@acu.edu.au

The Self-Compassion Construct

Self-compassion research has exponentially grown, as reflected in hundreds of studies and numerous recent meta-analyses (Ferrari et al., 2019; Inwood & Ferrari, 2018; Kirby et al., 2017; Phillips & Hine, 2021; Wakelin et al., 2022; Neff's (2003a, 2003b) SCS has been cited over 7,400 times (based on google scholar, January 2, 2023) and has had a defining influence on the conceptualization of self-compassion. Neff (2003a) defined self–compassion as a healthy attitude toward oneself that protects against poor mental health. This involves balancing increased compassion and reduced uncompassionate responses to life difficulties and personal suffering. More specifically, Neff's (2003a, 2003b, 2023) self-compassion model includes six core constructs.

The three positively oriented scales and example items are as follows:

- Self-Kindness (SK, e.g., "I try to be loving toward myself when I'm feeling emotional pain").
- Common Humanity (CH, e.g., "When things are going badly for me, I see the difficulties as part of life that everyone goes through").
- Mindfulness (MI, e.g., "When I'm feeling down I try to approach my feelings with curiosity and openness").

The three negatively oriented scales and example items are as follows:

- Self-Judgment (SJ, e.g., "I'm disapproving and judgmental about my own flaws and inadequacies").
- Isolation (IS, e.g., "When I think about my inadequacies it tends to make me feel more separate and cut off from the rest of the world").
- Overidentification (OI, e.g., "When something upsets me I get carried away with my feelings").

The three positively oriented scales are summed to form a global compassionate Self-Responding (CS) score. The sum of the three negatively oriented scales is typically reverse-scored to form a global reverse-scored uncompassionate self-responding (RUS) score. A total SCS score is the sum of CS and RUS. Neff and colleagues provide more detail on the theoretical and psychometric basis of the SCS in numerous publications (e.g., Neff, 2003a, 2003b, 2016, 2023; Neff et al., 2018, 2019). Although the use of both individual subscales and a total score are recommended, Kumlander et al. (2018) note that most research using the SCS relies almost exclusively on a total score.

Although long recognized as an essential construct in Eastern philosophies such as Buddhism, a recent surge in self-compassion research in clinical psychology is stimulated by its association with a plethora of mental health outcomes (Ferrari et al., 2019; Kirby et al., 2017). MacBeth and Gumley's (2012) meta-analysis reported that lower self-compassion was significantly related to psychopathology, including depression, anxiety, and stress symptoms. Kirby et al.'s (2017) meta-analysis reported that self-compassion interventions significantly affected self-compassion and mental health outcomes (depression, anxiety, and well-being). Ferrari et al.'s (2019) meta-analysis found that self-compassion interventions had beneficial effects across 11 psychosocial outcomes, including eating behavior, rumination, self-compassion, stress, depression, MI, self-criticism, and anxiety. Further meta-analyses reported that self-compassion benefits the psychological well-being of diverse populations, including people with multiple sclerosis (Simpson et al., 2023), sexual and gender minorities (Helminen et al., 2023), patients with breast cancer (Fan et al., 2023), Asian communities (Kariyawasam et al., 2023), those who are self-critical (Wakelin et al., 2022), and those with body image concerns (Turk & Waller, 2020).

The SCS Factor Structure Controversy

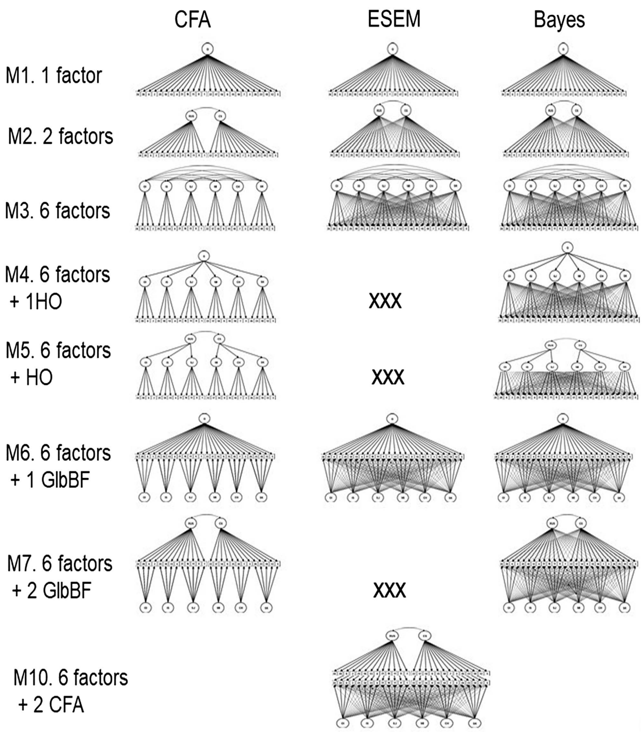
Self-compassion research is largely driven by Neff's (2003a, 2003b, 2023) six-factor SCS. Indeed, Ferrari et al. (2019) noted that the widespread use of the SCS has contributed to unifying the growing body of self-compassion research. Moreover, there is solid psychometric support for SCS's factor structure based on the six SCS factors. However, there is controversy about SCS's higher order structure—and implications for theory, scoring the instrument, and clinical application (Ferrari et al., 2022). In particular, the issue is whether global factors should be used to represent SCS responses. If so, should this be a single global factor or two global factors representing compassionate self-responding (CS) and RUS (Cleare et al., 2018; J. Costa et al., 2016; Muris & Petrocchi, 2017)?

Some argue that self-compassion is a bipolar continuum, like hot and cold (Neff, 2022, 2023). This view implies that self-criticism and self-compassion are inexorably connected and should be measured together as a single bipolar construct. In contrast, Muris and colleagues argue that negative RUS factors (e.g., self-criticism) are distinct from positive CS factors (self-kindness). They contend that using a single, bipolar global score contaminates the self-compassion construct with psychopathological characteristics (Muris, 2016; Muris & Otgaar, 2020; Muris & Petrocchi, 2017; Muris et al., 2019; Muris et al., 2021). Others (e.g., Brenner et al., 2017; J. Costa et al., 2016; Kumlander et al., 2018) argue that researchers should distinguish between positive and negative components of self-compassion but that both are essential components of the self-compassion construct.

Brenner et al. (2017) noted that SCS is typically represented as six specific factors and one higher order factor (see "6CFA + 1HO" in Figure 1). The authors examined the SCS structure using oblique, higher order, and bifactor models. Their results supported a bifactor model with two global factors ("6CFA + 2GlbBF" in Figure 1) over a bifactor model with a single global factor (see "CFA6: 6 + 1GlbBF" in Figure 1). In particular, none of the unidimensional, higher order, and bifactor models positing a global SCS factor fit the data well. In contrast, corresponding models positing two global factors always had a better fit than the corresponding model with only one global factor. The two-global-bifactor model and the oblique six-factor model best fit the data. Brenner et al. defended the bidimensionality representation in relation to theory (e.g., Gilbert et al., 2011). They concluded that SCS is best represented as six specific and two general factors. Brenner et al. argued that their research called into question previous research that confounded the effects of compassion's positive and negative components through reliance on a total score.

Figure 1

Selected Structural Diagrams of Confirmatory Factor Analysis (CFA), Exploratory Structural Equation Modeling (ESEM), and Bayes Models



Note. All models are based on responses to the 26 items from the Self-Compassion Scale. Models 1–7 (M1–M7) are the basic models that we fit with both CFA and Bayes (Models M4, M5, and M7 could not be fit with ESEM). M1–M3 are first-order models. M4 and M5 are higher order factor models, positing 1 or 2 higher order (HO) factors. M6 and M7 are bifactor models positing one or two global bifactors (GlbBF). M10 is idiosyncratic to ESEM used by Neff et al., 2019) as an alternative to a true bifactor model (that cannot be estimated with ESEM). ESEM models denoted by XXX cannot be estimated with ESEM as currently configured, although Marsh et al. (2014) proposed a two-step approach to test HO factor models referred to as ESEM-within-CFA.

Neff (2003a) initially argued for a six-factor first-order structure with a single higher order factor. However, based on evolving statistical modeling research using the Mplus statistical package (Morin, Arens, et al., 2016; Tóth-Király & Neff, 2021), Neff (2023) and Neff et al. (2019) subsequently rejected this SCS representation in favor of a more sophisticated bifactor exploratory structural equation model (BI-ESEM). They argued that this statistical approach combined the best features of EFA and CFA (Marsh et al., 2014; Morin, Arens, et al., 2016), was more suitable for testing multidimensional constructs, and was more consistent with the psychological theory underlying the SCS.

In pursuit of this goal, Neff et al. (2019) conducted the most extensive analysis of SCS's factor structure ever undertaken. Using a large sample (N = 11,685) representing 20 countries and 13 languages, Neff et al. compared various solutions based on CFA, ESEM, and BI-ESEM—a substantive-methodological synergy. Neff and colleagues argued that because of the ordinal nature of SCS responses (i.e., a 5-point Likert scale), weighted least square mean and variance adjusted (WLSMV) is a more appropriate estimator than the traditional maximum likelihood estimation (e.g., Bandalos, 2014, but also see Sass et al., 2014).

Neff et al. (2019; and replicated in our analyses) found that first-order factor analyses based on CFA and ESEM showed good support for SCS's six-factor structure (also see subsequent discussion of Table 1, for correlations among the six SCS factors). However, CFA models did not clearly support models with either a single or two global factors. The ESEM bifactor model with a single global factor fit the data well, and the global bifactor was well-defined. The ESEM model with two global CFA factors also fit the data well. However, both global factors were weak and poorly specified. On this basis, Neff and colleagues argued for using six subscale scores or a total SCS score. The total SCS score represented self-compassion as a bipolar construct that varied along a positive-to-negative continuum (i.e., a bipolar factor). However, based on the failure of the ESEM model with two global CFA factors, Neff et al. (2019) rejected the representation of compassion as separate positive and negative components. This research contributed substantially to the debate on whether negative compassion should be part of the compassion construct (e.g., Muris & Petrocchi, 2017). Comparing these two ESEMs with one and two global factors underpins the argument by Neff and colleagues for reliance on a single global factor. Thus, this methodological issue has critical implications for theory, research, practice, and policy.

Rakhimov et al. (2023) subsequently extended the analysis of Neff et al. (2019). Rakhimov et al. tested Neff et al.'s model, including CFAs, ESEMs, and bifactor models. They related the ESEMs with one and two global factors to positive and negative mental health indicators. Like Neff et al., Rakhimov et al. found that ESEMs with one or two global factors each fit the data well. However, only the ESEM with a single global bifactor was well defined. In the ESEM with two global factors, the global factors had weak, inconsistent factor loadings, and did not explain much of the variance. Rakhimov et al. (p. 11) concluded that their research supported "Neff's (2022) view that a total SCS score reflects an entire bipolar continuum and thus has more explanatory power than the two separate scores" and that

> using two global scores rather than one could lead to less accurate findings so that it is preferable to use one total SCS score instead. This is particularly important to consider when the SCS is being used in trials and clinical settings.

Interestingly, Brenner et al. (2017), Neff et al. (2019), and Rakhimov et al. (2023) all used many of the same models but came up with diametrically opposed conclusions about the appropriateness of the one- or two-global-factor representation of the SCS structure. The critical difference is that Brenner et al. relied on CFA models, whereas Neff et al. and Rakhimov et al. extended the models to include new and evolving ESEM models. Thus, Brenner et al. argued in favor of a bifactor CFA model with six specific and two global factors rather than the corresponding CFA bifactor model with only one global factor. In contrast, Neff et al. and Rakhimov et al. argued for a bifactor ESEM model with one global bifactor rather than an ESEM model with two global CFA factors. However, Neff et al. and Rakhimov et al. actually found that the ESEM model with two CFA global factors fit the data slightly

Table 1

Comparison of Alternative Models of Self-Compassion Factor Structure: Confirmatory Factor Analysis (CFA) or Exploratory Structural Equation Model (ESEM)

		Continuous estimator						Categorical estimator					
Model	*df*	Chi-sq	RMSEA	CFI	TLI	Corr	*SE*	Chi-sq	RMSEA	CFI	TLI	Corr	*SE*
CFA													
Model 1: CFA 1 + 0	299	35,063	.100	.691	.664			80,132	.151	.747	.725		
Model 2: CFA 2 + 0	298	13,880	.062	.879	.868	.60	.009	31,573	.095	.901	.892	.60	.006
Model 3: CFA 6 + 0	284	7,821	.048	.933	.923			17,463	.072	.946	.938		
Model 4: CFA 6 + 1HO	293	15,445	.067	.865	.851			54,103	.125	.830	.811		
Model 5: CFA 6 + 2HO	292	8,903	.050	.924	.915	.62	.01	21,391	.079	.933	.926	.64	.007
Model 6: CFA 6 + 1GlbBF[a]	273	12,817	.063	.889	.867			48,558	.123	.847	.818		
Model 7: CFA 6 + 2GlbBF[a]	272	5,346	.040	.955	.946	.63	.009	12,594	.063	.961	.953	.64	.007
ESEM													
Model 1: ESEM 1 + 0	299	35,063	.100	.691	.664			80,132	.151	.747	.725		
Model 2: ESEM 2 + 0	274	12,711	.062	.890	.869	.57	.009	38,131	.109	.88	.858	.56	.006
Model 3: ESEM 6 + 0	184	2,184	.030	.982	.969			4,748	.046	.986	.974		
Model 4: ESEM 6 + 1GlbBF[a]	164	1,576	.027	.987	.975			3,304	.040	.990	.980		
Additional ESEM models													
Model 8: ESEM 6 + 2ESEM1[a]	145	1,035	.023	.992	.982			2,268	.035	.993	.985		
Model 9: ESEM 6 + 2ESEM2	145	1,035	.023	.992	.982	.61	.009	2,268	.035	.993	.985	.62	.006
Model 10: ESEM 6 + 2CFA	157	1,155	.023	.991	.982	.20	.139	2,399	.035	.993	.985	.09	.054
Model 11: SetESEM	217	1,943	.026	.985	.977	.62	.009	3,531	.036	.990	.984	.62	.009

Note. Chi-sq = chi-square; *df* = degrees of freedom; CFI = comparative fit index; TLI = Tucker–Lewis Index; RMSEA = root-mean-square error of approximation. Model. Corr = correlation between positive and negative compassion (for those models that contain this estimate). *SE* = standard error. Model description (see Figure 1 and Table 2): the first number refers to the number of first-order factors and the second number refers to the number of global factors. The global factors are traditional higher order factors (HO), global bifactors (GLB), global ESEM factors estimated within the ESEM model (ESEM), or global CFA factors estimated separately from the ESEM model (CFA). Within the two 6 + 2ESEMs, rotation was either orthogonal (6 + 2ESEM1) or not orthogonal (6 + 2ESEM2). ESEM Models 5–7 (i.e., those corresponding to CFA Models 5–7) cannot be estimated with ESEM as currently configured. ESEM Models 8–11 are alternative models that we explored to counter these limitations in ESEM.
[a] Models that are true bifactor models.

better than the corresponding model with only one global factor. Nevertheless, they rejected the model because the two global factors were poorly defined, and the results did not make sense. Hence, the critical issue in this controversy is the ESEM model with two global CFA factors—why the global factors were ill-defined and why the results did not make sense.

The Present Investigation

The Need for Substantive-Methodological Synergy

In the present investigation, we reanalyze data from Neff et al. (2019); $N = 11{,}865$ responses from 20 SCS studies. In our reanalysis of Neff et al. (2019), we tested models considered by Neff et al., Rakhimov et al. (2023), Brenner et al. (2017), and a variety of new models. First, we critique the previously used models. Then, we posit new models to represent the SCS factor structure better and apply evolving Bayes models. This allows us to consider additional models that could not be tested with methodological approaches used in earlier studies. The end goal of this approach is to make a substantive contribution (what is the factor structure of the SCS) and a methodological conclusion (what are the best methods for analyzing measures with hierarchical factor structures).

Our study is a substantive-methodology synergy where we apply evolving statistical methodology to substantively important issues with implications for theory, policy, and clinical practice. In this approach, we take the role of skilled data detectives, following many alternative leads (Marsh et al., 1999). Like detectives, we develop appropriate tests of plausible counterinterpretations of critical conclusions, pursue these tests as part of an ongoing research program, and make a case for the most defensible interpretations.

In our analyses, we considered models described in Table 2 and illustrated in Figure 1. We broadly classify these as CFA, ESEM, and Bayes for present purposes. We begin with a set of seven CFA models:

- Three first-order models positing one, two, or six SCS factors (Models 1–3).
- Two higher (second-) order models positing one or two higher order (HO) factors (Models 4 and 5).
- Two bifactor models positing one or two global bifactor (GlbBF) factors (Models 6 and 7).

We then estimated this set of seven CFA models with ESEM (although not all the models can be fit with ESEM as currently configured) and Bayes. For both ESEM and Bayes, we also explore additional models that focus on issues specific to our study (described in Table 2).

Methodology and Sample

Because the sample, instrument, basic psychometric properties (e.g., reliability estimates of the six SCS scales), and statistical analyses are presented in greater detail by Neff et al. (2019; also see Brenner et al., 2017; Rakhimov et al., 2023), we summarize theses only briefly for present purposes. As described in more detail by Neff et al. (2019), our sample included 11,685 respondents

Table 2
Description of Basic and Alternative Models (Also See Figure 1 and Tables 1, 4 and 5)

Model	Basic models
Model 1	One specific first-order (FO) factor, zero global factors. The structure is the same for CFA, ESEM, and Bayes (i.e., all items load on one factor representing total SCS).
Model 2	Two specific factors, zero global factors. The CS (positively oriented) items load on a specific CS factor, and all RUS (negatively oriented) items loading on the specific RUS (reverse-scored uncompassionate) factor. For 2CFA + 0, items do not cross-load. For 2ESEM + 0 and 2Bayes + 0, all items cross-load on both factors.
Model 3	Six specific factors, zero global factors. Items from each of the six SCS factors load on the SCS factor it is designed to measure. For 6CFA + 0, items do not cross-load. For 6ESEM + 0 and 6Bayes + 0, all items are allowed to cross-load on all factors.
Model 4	Six specific factors, zero higher order (HO) factor. Each item loads on the SCS factor it is designed to measure. The six FO factors are constrained to be uncorrelated with each other and the HO factor. All six FO factors load on a single higher order factor. For 6CFA + 1HO, items do not cross-load. For 6Bayes + 1HO, all items are allowed to cross-load on all six FO factors. 6ESEM + 1HO cannot be estimated.
Model 5	Six specific factors, two higher order (HO) factors. Each item loads on the SCS factor it is designed to measure. The six FO factors are constrained to be uncorrelated with each other and the two HO factors. The three positive FO factors load on the higher order CS factor and the three negative FO factors load on the HO RUS factor. For 6CFA + 2HO, items do not cross-load. For 6Bayes + 2HO, all items are allowed to cross-load on all six FO factors and both HO factors. 6ESEM + 2HO cannot be estimated.
Model 6	Six specific factors, one global bifactor (GlbBF). Each item loads on the SCS factor it is designed to measure. The six FO factors are constrained to be uncorrelated with each other and the single global bifactor (GlbBF) factor. All items also load on the GlbBF. For 6CFA + 1GlbBF, items do not cross-load. For 6ESEM + 1GlbBF and 6Bayes + 1GlbBF, all items are allowed to cross-load on all six FO factors.
Model 7	Six specific factors, two global bifactor (GlbBF) factors. Each item loads on the SCS factor it is designed to measure. The six FO factors are constrained to be uncorrelated a with each other and the two GlbBF factors. The three positive FO factors load on the higher order CS factor and the three negative FO factors load on the higher order RUS factor. For 6CFA + 2GlbBF, items do not cross-load. For 6Bayes + 2GlbBF, all items are allowed to cross-load on all six FO factors. 6ESEM + 2GlbBF cannot be estimated.
	Alternative ESEM models
Model 8	6ESEM + 2ESEM1 is like Model 5 in that there are six specific factors and two global factors. However, all eight factors (six FO specific and two global) factors were specified to be uncorrelated. This is a true bifactor, but requiring the two higher order factors to be uncorrelated does not allow tests of this correlation. As currently configured ESEM does not allow some factors to be correlated and others uncorrelated.
Model 9	6ESEM + 2ESEM2 is like Model 8 (6ESEM + 2ESEM1) except all eight factors (six FO specific and two global) factors were specified to be correlated. Because all the factors are correlated, this is not a bifactor model. As currently configured ESEM does not allow some factors to be correlated and others uncorrelated.
Model 10	6ESEM + 2CFA combines an ESEM for the six FO factors, and a CFA for the two global factors. For the ESEM component, items from each of the six SCS factors load on the SCS factor it is designed to measure and all items are allowed to cross-load on all factors. For the CFA component, the CS (positively oriented) items load on a global specific CS factor and all RUS (negatively oriented) items loading on the global RUS factor; for the two global factors, items do not cross-load.
Model 11	6ESEM + 2ESEM3 is a SetESEM with all the positively oriented items forming one set and all the negatively worded items forming the other set. Within each set, we posited a bifactor model with one global factor (i.e., are three specific factors and one global factor). Within each set, all factors are uncorrelated. However, across the two sets factors from one set are allowed to be correlated with factors from the other set. This provides an estimate of the correlation between the two global factors. However, because the specific factors are correlated, it is not a true bifactor model.
	Alternative Bayes models
Models 2 A, 5 A, and 7 A are like Models 2, 5, and 7	Both set of models posit one global factor for positively oriented item (CS) and for negatively oriented items (RUS). However, in alternative models (2 A, 5 A, and 7 A) items do not cross-load across the global factors. For the original set of Models (2, 5, and 7) positively oriented items are allowed to cross-load on the RUS global factor, and the negatively oriented items are allowed to cross-load on the CS global factor. Results based on Models 2 A, 5 A, and 7 A are nearly the same as for Models 2, 5, and 7.

Note. See Figure 1 for a diagrammatic presentation of models and supplemental materials for Mplus syntax to fit Bayes models (Mplus syntax for CFA and ESEM models is presented by Neff et al., 2019). CFA = confirmatory factor analysis; ESEM = Exploratory Structural Equation Model; SCS = Self-Compassion Scale; CS = compassionate self-responding; RUS = reverse-scored uncompassionate self-responding.

(3,296 males, 8,367 females, 22 unspecified), aged between 18 and 83 ($M = 32.29$, $SD = 8.28$). Participants came from 20 international samples—including 12 translations. All respondents completed the 26-item SCS self-report instrument using a 5-point Likert response scale. In addition, Neff et al. (2019) present a summary of basic psychometric properties (e.g., coefficient α estimates of reliability).

To maintain comparability, we began using many of the statistical analyses and models used by Neff et al. (2019) and Rakhimov et al. (2023). They fit a wide variety of different CFA, ESEM, and BI-ESEM models. They argued that the weighted least squares mean- and variance-adjusted estimator (WLSMV) was more appropriate for the SCS's 5-point Likert response scale. Following Marsh and colleagues (e.g., Marsh et al., 2014; Morin, Arens, et al., 2016), they argued for the application of ESEM that combined the best features of EFA and CFA that were more suitable for testing multidimensional constructs and was more consistent with the psychological theory underlying the SCS. Morin, Arens, et al. (2016), Neff et al. (2019), and Rakhimov et al. (2023) argued that bifactor models provide a better approach to testing higher order factor structures than traditional higher order (HO) factor structures. They note that HO factors are based on covariation between first-order factors rather than items that define the first-order factor structures. In evaluating models, they relied substantially on traditional fit indices and accepted guidelines of fit (Hu & Bentler, 1999; Marsh et al., 2014); the comparative fit index (CFI; .95 is good, 90 is acceptable), the Tucker–Lewis Index (TLI; .95 is good, .90 is acceptable), and the root-mean-square error of approximation (RMSEA; .06 is good, .08 is acceptable). However, following Marsh et al. (2014) and others, they emphasized that the interpretation of the appropriateness of a model should not be based solely on goodness of fit.

Neff (2016, 2023) and Neff et al. (2019) contended that a bifactor model provides a better theoretical fit with her conceptualization of self-compassion than a higher order model. She argued that behaviors assessed by individual items directly represent self-compassion as a general construct in addition to its constituent group components. Coupled with a preference for ESEM over CFA, Neff et al. (2019) argued for the appropriateness of a BI-ESEM framework (Morin, Arens, et al., 2016; Morin, Boudrias, et al., 2016; Tóth-Király & Neff, 2021). Thus, Morin et al. (2020) argue for the need to systematically compare solutions based on CFA, ESEM, bifactor-CFA (BI-CFA), and bifactor ESEM (BI-ESEM). In particular, Morin et al. (2020) contend that researchers should use ESEM for multidimensional constructs, bifactor models for hierarchically related constructs, and BI-ESEM when applications include both sources of construct-relevant psychometric multidimensionality (see Morin et al., 2020, for further discussion of bifactor models, the difference between CFA and ESEM, and BI-ESEM framework).

Following this BI-ESEM approach, Neff et al.'s (2019) critical comparisons should be between ESEM models position the six SCS first-order specific factors in combination with no global factors (6ESEM + 0GlbBF), one global bifactor (6ESEM + 1GlbBF), two global bifactors (6ESEM + 2GlbBF). 6ESEM + 0GlbBF and 6ESEM + 1GlbBF are straightforward ESEM models. However, as presently conceptualized, Mplus cannot test the 6ESEM + 2GlbBF structure (see "6Bayes + 2GlbBF" in Figure 1 and Table 2 that has the same structure as 6ESEM + 2GlbBF if it could be estimated). Recognizing this limitation in ESEM, Morin et al. (2020) and Tóth-Király et al. (2018) proposed logical variations of 6ESEM + 2GlbBF to circumvent this problem and recommended 6ESEM + 2CFA (see Model 10 in Figure 1 and Table 2) as the most appropriate approximation of a true 6ESEM + 2GlbBF. A major focus of our study is evaluating the appropriateness of this model as an approximation of the true 6ESEM + 2GlbBF. This reanalysis is critical because Neff et al. recommended use of a total SCS score rather than separate positive (CS) and negative (RUS) SCS scores is based on the comparison of ESEM models with six specific ESEM factors and either one global bifactor (6ESEM + 1GlbBF in Figure 1 and Table 2) or two global CFA factors (6ESEM + 2CFA, see Figure 1 and Table 2) rather than the more appropriate 6ESEM + 2GlbBF that cannot be tested with ESEM.

Problems With the 6ESEM + 2CFA Model

6ESEM + 2CFA Is Neither an ESEM Nor a Bifactor Model

A fundamental problem with 6ESEM + 2CFA is that it is neither an ESEM nor a bifactor model. 6ESEM + 2CFA is not an ESEM model because items included in the ESEM are modeled as CFA factors outside the ESEM model. Thus, it is not part of the rotation that defines the ESEM factors. This is quite different from the bifactor ESEM model with one global factor where the global factor is part of the ESEM rotation. It is unclear, a priori, how important this distinction is—but it turns out to be critical. ESEM + 2CFA is not a bifactor model in that not all the items load on the global factors. So, in this sense, neither of the "global" factors is a global bifactor factor. Again, it is unclear, a priori, how important this distinction is.

Furthermore, as currently configured, Mplus cannot fit a true BI-ESEM in which the two global bifactors correlate with each other but are uncorrelated with the six specific (i.e., 6ESEM + 2GlbBF model). However, various approximations of the 6ESEM + 2GlbBF model are possible (see subsequent discussion), including the 6ESEM + 2CFA model used by Neff et al. (2019) and Rakhimov et al. (2023).

The 6ESEM + 2CFA model is neither a "true" ESEM model, nor a "true" bifactor model. Nevertheless, it has much of the logic of a bifactor model. Therefore, it might represent a reasonable compromise if it fits the data, the factors are well-defined, and the parameter estimates make sense. However, the two global factors were not well-defined, as both Neff et al. (2019) and Rakhimov et al. (2023) emphasized.

6ESEM + 2CFA Results Are Illogical and Internally Inconsistent

From our perspective as data detectives, the 6ESEM + 2CFA results are internally inconsistent and illogical compared to 6ESEM + 1GlbBF results. Thus, Neff et al. (2019), Rakhimov et al. (2023), and our reanalyses, all demonstrated that with the 6ESEM + 1GlbBF model, the one global factor is well-defined, explained much of the variance in SCS responses, and that all SCS items loaded substantially on it. 6ESEM + 1GlbBF represents a bipolar representation of self-compassion which assumes that CS and RUS are almost perfectly correlated (a correlation approaching 1.0 as RUS is reversed-scored RUS). To the extent that the correlation between global CS and RUS factors is significantly (and meaningfully) less than 1, then models with two global factors are supported. Hence, the critical issue is how close to 1 the correlation between the two global factors in 6ESEM + 2CFA.

If these two ESEM models were comparable, the 6ESEM + 1GlbBF with one global factor would be nested under the 6ESEM + 2CFA model with two global factors. This would facilitate the comparison of the two models in terms of goodness of fit and parameter estimates. In particular, if the fit was similar and the correlation between the two global factors approached 1.0, one might argue for the more parsimonious model with one global factor.

However, it makes no sense that one global factor (6ESEM + 1GlbBF) is well-defined and able to explain so much of the variance, but two global factors (6ESEM + 2CFA) explain almost none of the variance. If there truly is only one coherent global factor, then the two-factor solution would represent two halves of this global factor. Therefore, the two global factors should explain as much variance as one global factor, and the two factors should be almost perfectly correlated. In this case, the main criticism of the two-factor solution would be that it lacks parsimony, not that it explains less variance. However, in the 6ESEM + 2CFA model, the empirically estimated correlation between the two global factors is close to zero and not statistically significant in Neff (2003b), Rakhimov et al. (2023), or our reanalysis. Mark Twain once noted that when the clock strikes 13, ignore the first 12. Obviously, there is a problem with the 6ESEM + 2CFA model that calls interpretations based on it into question. Because this is a logical issue not specifically related to the 6ESEM + 2CFA model, it is surprising that it has not been identified in peer reviews of the many SCS publications making this claim.

Use of a Bayes Estimator

SEM models can be estimated within the traditional Bayesian framework (Gelman et al., 2004). Bayes uses small variance priors to relax the SEM model to accommodate minor differences between the model and the observed data (Asparouhov & Muthén, 2021; Marsh, Guo, et al., 2020; Muthén & Asparouhov, 2012; also see Mplus syntax in supplemental materials). The rationale for this approach is like the logic of ESEMs used by Neff et al. (2019). Indeed, there is an ongoing debate about whether ESEM with target rotation is as good as a Bayes model with appropriately defined priors. Guo et al. (2019) compared the two approaches using real and simulated data. Mostly the results were similar. Bayes did marginally better, but ESEM also performed well. ESEM is easier for applied researchers to understand, but Bayes is more flexible. In particular, Bayes allows us to fit a true "BI-ESEM-like" model with two global factors (i.e., 6Bayes + 2GlbBF in Figure 1 and Table 2) and a 6Bayes + 1GlbBF model. Like ESEM, all the "nontarget loadings" (i.e., the loadings hypothesized to be small) were hypothesized to be zero and have small priors; that is, $\sim N(0, .01)$ that is similar to the logic of the ESEM approach; see Asparouhov and Muthén (2009, 2021) and Guo et al. (2019). Every SCS item was allowed to load on all six specific + 2 global factors. However, the Bayes approach allowed us to test the actual (6Bayes + 2GlbBF) model with the two global correlated factors and six specific factors that were uncorrelated with each other and with the two global factors. As noted earlier, this model cannot be tested with BI-ESEM as currently operationalized in Mplus. Asparouhov and Muthén (2021) described recent advances in Bayesian model fit for evaluating SEMs (BSEM) and building well-fitting BSEMs. These included new Bayesian adaptations of the approximate fit indices RMSEA, CFI, and TLI, as well as the Bayesian adaptation of the Wald test for nested models when sample sizes are substantial (as in our study). Although currently only available for Bayes models based on continuous data, this is an area of ongoing development. They emphasized that these new features were effective with simulated data and could "enlighten real data applications"—a focus of the present investigation.

Research Hypotheses and Questions

Recent research has shown strong support for the 6ESEM + 1GlbBF model and a strong lack of support for the 6ESEM + 2CFA (Neff et al., 2019; Rakhimov et al., 2023). This evidence has been the primary basis for substantially important interpretations with critical policy and clinical practice implications. Our overarching research aim is to demonstrate that the interpretation of the 6ESEM + 2CFA is seriously flawed, casting doubt on recommendations based on it. In testing this overarching hypothesis, we explore a variety of alternative models. In this sense, the article makes a statistical contribution (how to deal with the problem of a BI-ESEM model with two global factors) and a substantive contribution to self-compassion research. Furthermore, the issues we address with SCS responses are evident in many clinical measurement instruments (see annotated bibliography in supplemental material Section 1). From this perspective, our study is a substantive-methodological synergy (Marsh & Hau, 2007), applying evolving statistical practice to substantially important issues with critical implications for theory, policy, and clinical practice. In pursuit of this overarching aim, we offer the following research hypotheses:

1. A comparison of alternative (CFA, ESEM, BI-ESEM, and Bayes) models of SCS responses will demonstrate flaws in interpreting the 6ESEM + 2CFA model (see earlier discussion). This is critical because this model is central in recent self-compassion research, debates on a bipolar representation of self-compassion, and claims for the superiority of the 6ESEM + 1GlbBF model.

2. For all CFA, ESEM, BI-ESEM, and Bayes analyses (except 6ESEM + 2CFA), models with the two overarching factors representing positive CS and negative RUS components of self-compassion will fit the data better than corresponding models that collapse these two factors into a single overarching factor.

3. For all CFA, ESEM, BI-ESEM, and Bayes models (except 6ESEM + 2CFA), the estimated correlation between two overarching factors representing positive CS and negative RUS components of self-compassion will be substantial (e.g., .6–.7 as reported in related research, e.g., Neff, 2003b; Neff et al., 2019). However, the correlation will be substantially less than 1.0 (that would support collapsing the two factors into a single overarching self-compassion factor).

Method

Participants and Measures

Neff et al.'s (2019) Data

The Neff et al.'s final sample included 11,685 respondents (3,296 males, 8,367 females, 22 unspecified) aged between 18 and 83 (M = 32.29, SD = 8.28). Participants came from 20 international samples (see Neff et al., for more detail). All respondents completed SCS, a 26-item self-report instrument—including 12 translations. Participants responded using a 5-point Likert response scale. Neff et al. summarize basic psychometric properties (e.g., reliability estimates). Neff et al. did analyses using Mplus 7.4 (Muthén & Muthén, 1998–2018) with the WLSMV that they argued was more appropriate for the SCS's 5-point Likert response scale. They fit a wide variety of different CFA, ESEM, and BI-ESEM models. In evaluating models, they relied substantially on traditional fit indices and accepted guidelines of fit (Hu & Bentler, 1999; Marsh et al., 2004)—the CFI (.95 for good, 90 for acceptable), the TLI (95 is good, .90 is acceptable), and the RMSEA (.06 is good, .08 is acceptable). However, following Marsh et al. and others, they emphasized that the interpretation of the appropriateness of a model should not be based solely on goodness of fit. Reanalysis of these data in the present investigation not only included all 11,685 cases and considered many of the same models but also introduced new Bayes models.

This study was based on a reanalysis of data from Neff et al. (2019) that was graciously provided by the Kristin Neff. Enquiries about access to the data should be sent to Kristin Neff (kristin@self-compassion.org). In the present investigation, code for analyses is provided by Neff et al. (2019) and in supplemental materials. The present investigation was not preregistered. Because the study was a reanalysis of secondary, anonymous data previously published in Psychological Assessment (Neff et al., 2019) and generously provided by the lead author of that article, it was deemed not to require further ethical approval.

Results

Goodness of Fit (Tables 1 and 3)

We begin with an overview of the goodness of fit (see Tables 1 and 3) for alternative models (Hypothesis 1) that was the primary focus of Neff et al. (2019) and Rakhimov et al. (2023) studies. In these analyses, we applied six estimation approaches (CFA, ESEM, and Bayes models with categorical and continuous estimators). Within each approach, we considered various models (see earlier discussion of Table 2 and Figure 1), some of which are idiosyncratic to particular approaches. The comparison of models positing overarching factors reflecting positive and negative compassion as one or two factors is particularly relevant. The key question is whether there is support for the contentions by Neff et al. and Rakhimov et al. that models with one overarching (positive CS + negative RUS) provide the best representation of these data or our a priori hypotheses that models with two overarching factors are needed (i.e., separate positive CS and negative RUS factors).

For all six approaches (CFA, ESEM, Bayes with categorical and continuous estimators), the two-factor (CS and RUS) fit substantially better than a one-factor model. However, neither of these parsimonious models provided an acceptable fit. In contrast, the fit of all six-factor models was reasonable. The fit was marginal for CFA models (CFIs ≥ .933) but improved substantially for ESEM (CFIs ≥ .982) and Bayes (CFI = .983) models.

For traditional higher order models, the CFA model with 1 HO factor provided a poor fit to the data (CFIs ≤ .90). In contrast, the fit if the CFA-2HO model was marginal (CFIs ≥ .92). Furthermore, Bayes models with one and two HO factors had similar fits that were consistently good (all CFIs = .983).

True bifactor models are those with an "a" in Tables 1, 3 and 4 BI-CFA models with two global factors had an acceptable fit (CFIs ≥ .955), whereas BI-CFA models with one global factor did not (CFIs ≤ .889). ESEM models with one (6ESEM + 2GlbBF) and two (6ESEM + 2CFA) global factors both provided excellent fits to the data (CFIs ≥ .987), but 6ESEM + 2CFA fit marginally better. Similarly, for the Bayes models, bifactor models with one and two global factors all fit the data well (CFIs ≥ .988), but the fit for the model with two global factors was marginally better. Additional ESEM models positing two global CFA factors are not true BI-ESEM models but embrace various aspects of the bifactor logic. These models all fit the data very well (CFIs ≥ .977). More broadly, models with two overarching factors (see Table 2 and Figure 1) fit the data better than those for one in most of the relevant comparisons:

- CFA: one versus two factors; one versus two HO factors; one versus two bifactors.

Table 3
Comparison of Alternative Models of Self-Compassion Factor Structure Using Bayes Estimator

	Continuous								Categorical				
Model	free	Chi-OBS	Chi-Rep	RMSEA	CFI	TLI	Corr	*SE*	free	Chi-OBS	Chi-Rep	Corr	*SE*
Model 1: BAYES 1 + 0	78	45,705	45,822	.114	.696	.67			130	34,831	35,541		
Bayes2: 2 2 + 0 (YesXLd)	105	16,367	16,489	.072	.891	.871	.58	.04	157	11,631	12,143	.60	.03
Model 3: BAYES 6 + 0	223	2,536	2,675	.032	.983	.974			275	1,649	1,894		
Model 4: BAYES 6 + 1HO	214	2,524	2,657	.032	.983	.975			266	1,608	1,848		
Model 5: BAYES 6 + 2HO	221	2,525	2,658	.031	.983	.975	.67	.122	273	1,603	1,839	.64	.113
Model 6: BAYES 6 + 1GlbBF[a]	249	1,713	1853	.008	.988	.998			301	1,088	1,302		
Model 7: BAYES 6 + 2 GlbBF[a]	261	1,082	1,221	.010	.992	.998	.62	.044	313	677	895	.63	.028
Alternative Bayes models													
Model 2 A: BAYES 2 + 0	79	17,794	17,926	.072	.881	.871	.60	.007	131	12,857	13,359	.61	.007
Model 5 A: BAYES 6 + 2HO	215	2,522	2,655	.031	.983	.975	.66	.117	267	1,605	1,841	.66	.099
Model 7 A: BAYES 6 + 2 GlbBF[a]	235	1,274	1,411	.008	.991	.998	.65	.01	287	838	1,315	.67	.008

Note. Chi-sq = chi-square (observed and replicated); free = number of free parameters; CFI = comparative fit index; TLI = Tucker–Lewis Index; RMSEA = root-mean-square error of approximation. Model. Corr = correlation between positive and negative compassion (for those models that contain this estimate). *SE* = standard error. Model description (see Figure 1 and Table 2): The first number refers to the number of first-order factors, and the second number refers to the number of global factors. The global factors are traditional higher order factors (HO) or global bifactors (GlbBF). For models with two global factors (Models 2, 5, and 7), we fit alternative models in which each item loaded on only one of the two global factors (also see supplemental materials Section 3 for Mplus syntax).
[a] Models that are true bifactor models.

Table 4
Correlations Among the Six Self-Compassion Scale (SCS) Factors in the Different Models

Model	Factor	Continuous						Categorical					
		SK	CH	MI	SJ	IS	OI	SK	CH	MI	SJ	IS	OI
CFA	Self kindness	—						—					
	Common Humanity	.74	—					.73	—				
	Mindfulness	.84	.82	—				.84	.81	—			
	Self judgment	.64	.38	.48	—			.66	.38	.49	—		
	Isolation	.55	.42	.53	.85	—		.56	.42	.53	.84	—	
	Over identification	.57	.43	.57	.91	.90	—	.58	.43	.59	.91	.83	—
ESEM	Self kindness	—						—					
	Common humanity	.64	—					.64	—				
	Mindfulness	.56	.60	—				.54	.59	—			
	Self judgment	.42	.23	.22	—			.40	.23	.21	—		
	Isolation	.48	.39	.34	.60	—		.48	.39	.33	.59	—	
	Over identification	.32	.26	.16	.46	.66	—	.34	.28	.16	.43	.66	—
Bayes	Self kindness	—						—					
	Common Humanity	.75	—					.74	—				
	Mindfulness	.59	.67	—				.69	.71	—			
	Self judgment	.67	.45	.19	—			.63	.41	.30	—		
	Isolation	.69	.57	.30	.77	—		.65	.52	.42	.75	—	
	Over identification	.57	.46	.15	.71	.82	—	.55	.43	.29	.72	.82	—

Note. Grey shaded coefficients refer to the correlation between specific CS and RUS factors Estimated correlations among the six self-compassion factors based on first-order models (see "6 + 0" models in Tables 1, 3 and 4). SK = self kindness; CH = common Humanity; MI = mindfulness; SJ = self judgment; IS = isolation; OI = over identification; CFA = confirmatory factor analysis; ESEM = exploratory structural equation model.

- ESEM: one versus two factors;
- Bayes: one versus two factors; one versus two HO factors; one versus two bifactors

However, particularly for the Bayes models, the differences in fit were not substantial.

Relations Among Self-Compassion Factors: Six Specific Factors and One or Two Global Factors (Table 4)

Correlations Among Six First-Order SCS Factors

The pattern of relations among factors is a critical feature given insufficient attention by Neff et al. (2019) and Rakhimov et al. (2023). The clinical usefulness of the six SCS factors requires that they are sufficiently distinct to be distinguishable. Support for the separation of CS and RUS compassion components requires that correlations among the three positive CS factors, and among three negative RUS factors, are higher than correlations between CS and RUS factors. Furthermore, implicit in the SCS design is the assumption that there are three distinct content areas, each represented by a positive CS and a negative RUS factor (self-kindness and self-judgment; CH and IS; and MI and overidentification).

We begin by evaluating correlations among six SCS factors (Table 4) with no global CS and RUS factors. Across the different sets of models (CFA, ESEM, Bayes with categorical and continuous estimators), the pattern of correlations is reasonably consistent. Correlations are consistently high among the three positive CS compassion factors (.56–.84; $Md\ r = .72$) and among the three negative RUS compassion factors (.46–.91; $Md\ r = .76$). Correlations between CS and RUS factors are also substantial, but mostly lower in size (.15–.66; $Md\ r = .42$). Correlations tend to be higher for CFA solutions and lower for ESEM solutions. Although there is an implicit matching between CS and RUS factors based on theory, this is not evident in the pattern of correlations. Thus, for example, IS is paired with CH (Neff, 2003a, 2003b), but IS is more highly correlated with self-kindness and MI than with CH.

Correlations Between Overarching CS and RUS

The critical feature for models positing overarching positive CS and negative RUS factors is the correlation between the two. Support for models with one overarching factor requires that the correlation between overarching CS and RUS factors approaches 1.0 (keeping in mind that typical practice is to reverse-score RUS so that it reflects a lack of uncompassionate responding). This would mean no information is lost by combining the two factors, supporting a model with one overarching factor based on parsimony.

Many models estimate the correlation between positive and negative compassion factors (Tables 1, 3 and 4). With two major exceptions, these are consistently around .6 for all the different approaches: CFA (.60−.64); ESEM (.56−.62); Bayes (.58−.67). Although not all these models are true bifactor models, the pattern of results is consistent across all the models. Thus, CS and RUS compassion factors are highly correlated but never sufficiently correlated to justify collapsing the two factors into a single factor.

We now discuss the two major exceptions—ESEM models stemming from the fact that a true BI-ESEM with two global factors cannot be estimated with ESEM as currently configured. 6 + 2ESEM1 (Tables 1, 3 and 4) model is a true BI-ESEM with two global factors. However, ESEM restrictions require all global and specific factors to be all uncorrelated (i.e., a true BI-ESEM) or all correlated (not a bifactor model). Hence, by definition, the correlation between the global CS and RUS compassion factors is necessarily 0 for 6 + 2ESEM1. Thus, although this is a true BI-ESEM and fits the data well, the zero correlation is an unrealistic limitation. Interestingly, all the zero correlations in this model are achieved through rotation rather than explicit constraint. Hence, 6ESEM + 2ESEM2 (with all factors correlated) has the same *df* and fit as 6 + 2ESEM1 (with all factors uncorrelated). Of course, the 6 + 2ESEM2 model is not a bifactor model. Nevertheless, the correlation between the two global factors (*r*s = .61 and .62) is similar to the correlations based on other models—including the true bifactor models.

6ESEM + 2CFA is critical to the present investigation. Although it is not a true BI-ESEM, its logic is consistent with a BI-ESEM, and it fits the data well. However, the estimated correlation between the CS and RUS CFA factors was small. Indeed, despite the large sample size, the correlation was not statistically significant and had a large *SE* ($r = .20$, $SE = .139$ for continuous data, $r = .09$, $SE = .054$ for categorical data). This finding substantially undermines support for this model. Indeed, if the correlations really were close to zero, there would be no justification for collapsing the two factors into a single (CS + RUS) compassion factor, as posited by 6ESEM + 1ESEM that Neff et al. (2019) and Rakhimov et al. (2023) favored.

How Well-Defined Are the Global Factors in Bifactor Models (Table 5)

The juxtaposition of the patterns of factor loadings relating items to global bifactors in the 6ESEM + 1GlbBF and the 6ESEM + 2CFA solutions is a critical feature (Table 5). Unless the global factors from both solutions are well-defined, interpretations based on their comparison are dubious. Well-defined global factors should have consistently high factor loadings across the different SCS items posited to define each global factor (i.e., target loadings). Clearly, the two global factors in the 6ESEM + 2CFA solution were not well-defined. Neff et al. (2019) and Rakhimov et al. (2023) juxtaposed the strength of the global factors in both models. Nevertheless, they gave insufficient attention to the fact that it is internally inconsistent to have a single global factor that is well-defined but for neither of the two global factors to be well-defined. This inconsistency has not received deserved scrutiny. Hence, the critical issue is the juxtaposition between the interpretations based on these two models (but also their inconsistency with interpretations of the corresponding CFA models). Thus, when the metaphorical clock strikes 13, a good data detective must look further to determine what went wrong.

Except for 6ESEM + 2CFA, all the global factors are well-defined for all the CFA, ESEM, and Bayes solutions (Table 5). As a result, all the target loadings designed to define each global factor are substantial. For solutions with one global factor (Table 5), factor loadings are consistently large for all CFA, ESEM, and Bayes solutions (mean factor loadings vary from .54 to .62). For factors with two global factors, target factor loadings (i.e., CS items on the CS global factor, RUS items on the RUS global factor) are even larger (mean factor loadings varying from .61 to .81).

The one exception is the critical 6ESEM + 2CFA solution, where the global factors are poorly defined (also see bifactor indices in supplemental material Section 2). For continuous data, 8 of 26 factor loadings are nonsignificant. For the positive global factor, the mean factor loading is .40. Furthermore, the factor loadings are not consistent over SCS components, and the global CS factor primarily reflects self-kindness. The mean factor loading for the RUS global factor is .23—reflecting primarily self-isolation. For categorical data, 10 of 26 factor loadings are nonsignificant. For the CS global factor, the mean factor loading is .20—again reflecting primarily self-kindness. For the negative global factor, the mean factor loading is .14—again reflecting primarily self-isolation.

Discussion

Positive CS and Negative RUS Self-Compassion Factors

All models positing six SCS factors fit the data at least reasonably well. However, the fit was improved substantially for ESEM and Bayes and marginally by treating responses as categorical rather than continuous. In addition, bifactor models fit the data better than models without global factors, and two-global-factor models fit the data at least marginally better than one-global-factor models.

The correlations among factors were consistent with the distinctiveness of CS and RUS components of compassion. Correlations within the three CS factors and within the three RUS factors were higher than correlations between CS and RUS factors. The estimated correlation between the global CS and RUS factors (consistently around .6 except for the 6ESEM + 2CFA model) was substantial but substantially less than 1.0. Except for the 6ESEM + 2CFA model, the global factors were well-defined.

Bifactor models with two global factors were well-defined by CFA and Bayes, but the current versions of ESEM cannot test 6ESEM + 2GlbBF. We explored several approximations of the BI-ESEM but none were satisfactory. Particularly problematic is the 6ESEM + 2CFA model used by Neff et al. (2019). This is important as this model is the focus of much applied research and is critical in recent self-compassion discussions. Although 6ESEM + 2CFA fit the data well, its results were illogical. Problems with this model are clearly evident in its poorly defined global factors. In contrast to other models resulting in correlations around .6 between CS and RUS self-compassion factors, 6ESEM + 2CFA estimated the correlation to be close to zero. These results substantially undermine valid interpretations based on this model and clearly do not justify collapsing the two factors into a single (CS + RUS) compassion factor. Thus, the current results do not support the interpretation and application of the total SCS score in IS.

The 6ESEM + 2CFA results are illogical compared to the 6ESEM + 1ESEM. Logically a model with one global factor should be nested under a corresponding model with two global factors. 6ESEM + 1ESEM fits the data very well and has a well-defined global factor. Logically, 6ESEM + 2ESEM must also have two well-defined global factors that are highly correlated. Of course, we could not actually fit 6ESEM + 2ESEM and tried to approximate it with 6ESEM + 2CFA. However, the 6ESEM + 2CFA results are illogical concerning logical expectations for an appropriate 6ESEM + 2ESEM. Critically the logical pattern of results we expected was evident in the bifactor-Bayes models, where the appropriate models with structures like the 6ESEM + 2GlbBF could be estimated. Indeed, the logical pattern was also evident in CFA models, even though the fit was not as good. Based on our results, we reject the 6ESEM + 2CFA results as implausible.

Table 5

Factor Loadings Relating the 26 Self-Compassion Items (SCS) to Global Factors in Selected Confirmatory Factor Analysis (CFA), Exploratory Structural Equation Modeling (ESEM), and Bayes Models

Self Compassion Scale (SCS) items	CFA (Continuous)			CFA (Categorical)			ESEM (Continuous)								
	6 + 1	6 + 2		6 + 1	6 + 2		6 + 1	Orthogonal 6 + 2ESEM		Correlated 6 + 2ESEM		6 + 2CFA		SET ESEM	
	GLB	CS	RUS	GLB	CS	RUS	GLB	CS	RUS	CS	RUS	CS	RUS	CS	RUS
SCSK5	**.46**	**.69**		**.59**	**.73**		**.55**	**.65**	.20	**.76**	−.03	**.58**		**.69**	
SCSK12	**.54**	**.74**		**.66**	**.80**		**.61**	**.68**	.27	**.81**	.03	**.64**		**.76**	
SCSK19	**.55**	**.76**		**.67**	**.82**		**.64**	**.71**	.27	**.81**	.01	**.62**		**.78**	
SCSK23	**.57**	**.65**		**.66**	**.74**		**.70**	**.63**	.35	**.63**	.08	**.38**		**.73**	
SCSK26	**.55**	**.72**		**.67**	**.77**		**.71**	**.73**	.27	**.77**	−.04	**.47**		**.76**	
SCCH3	**.67**	**.57**		**.65**	**.59**		**.44**	**.51**	.13	**.52**	.01	**.29**		**.51**	
SCCH7	**.64**	**.51**		**.62**	**.54**		**.35**	**.48**	.03	**.50**	−.02	**.29**		**.46**	
SCCH10	**.65**	**.58**		**.64**	**.62**		**.42**	**.54**	.10	**.57**	.02	**.34**		**.53**	
SCCH15	**.74**	**.68**		**.72**	**.70**		**.56**	**.63**	.18	**.64**	.00	**.33**		**.64**	
SCMI9	**.63**	**.58**		**.63**	**.62**		**.48**	**.53**	.16	**.54**	−.01	**.26**		**.53**	
SCMI14	**.38**	**.70**		**.51**	**.74**		**.57**	**.64**	.19	**.64**	.01	**.26**		**.65**	
SCMI17	**.27**	**.68**		**.38**	**.72**		**.59**	**.62**	.22	**.63**	.02	**.30**		**.65**	
SCMI22	**.35**	**.66**		**.48**	**.67**		**.53**	**.63**	.16	**.71**	−.07	**.44**		**.64**	
SCSJ1	**.47**		**.70**	**.63**		**.74**	**.63**	.20	**.74**	.00	**.76**		**−.05**		**.77**
SCSJ8	**.70**		**.68**	**.69**		**.71**	**.60**	.18	**.66**	−.04	**.70**		**.19**		**.69**
SCSJ11	**.65**		**.67**	**.64**		**.72**	**.65**	.27	**.63**	.06	**.61**		**.14**		**.69**
SCSJ16	**.58**		**.77**	**.56**		**.81**	**.70**	.24	**.74**	.01	**.76**		**.15**		**.79**
SCSJ21	**.69**		**.65**	**.69**		**.71**	**.60**	.23	**.59**	.01	**.62**		**.37**		**.64**
SCIS4	**.41**		**.71**	**.53**		**.74**	**.62**	.20	**.66**	−.01	**.70**		**.22**		**.69**
SCIS13	**.49**		**.68**	**.65**		**.74**	**.61**	.21	**.60**	.02	**.62**		**.49**		**.63**
SCIS18	**.50**		**.62**	**.64**		**.67**	**.54**	.16	**.54**	−.02	**.59**		**.46**		**.57**
SCIS25	**.44**		**.71**	**.61**		**.74**	**.61**	.21	**.64**	.00	**.69**		**.28**		**.67**
SCOI2	**.75**		**.76**	**.75**		**.79**	**.65**	.20	**.75**	.00	**.80**		**.07**		**.77**
SCOI6	**.72**		**.72**	**.71**		**.75**	**.63**	.20	**.70**	−.01	**.74**		**.11**		**.73**
SCOI20	**.58**		**.61**	**.58**		**.65**	**.55**	.19	**.55**	.00	**.56**		**.26**		**.57**
SCOI24	**.60**		**.61**	**.60**		**.67**	**.56**	.22	**.53**	.05	**.54**		**.26**		**.57**
M Target	**.56**	**.66**	**.68**	**.62**	**.70**	**.73**	**.58**	**.61**	**.64**	**.33**	**.67**	**.40**	**.23**	**.64**	**.68**

Self Compassion Scale (SCS) items	ESEM (Categorical)									Bayes (Continuous)					Bayes (Categorical)				
	Orthogonal			Correlated															
	6 + 1	6 + 2ESEM		6 + 2ESEM		6 + 2CFA		SET ESEM		6 + 1	6 + 2		6 + 2		6 + 1	6 + 2		6 + 2	
	GLB	CS	RUS	CS	RUS	CS	RUS	CS	RUS	GLB	CS	RUS	CS	RUS	GLB	CS	RUS	CS	RUS
SCSK5	**.58**	**.67**	.22	**.79**	−.03	**.42**		**.73**		**.75**	**.82**		**.70**	−.02	**.69**	**.73**		**.74**	−.03
SCSK12	**.64**	**.71**	.29	**.83**	.03	**.47**		**.82**		**.70**	**.91**		**.72**	.04	**.65**	**.82**		**.77**	.03
SCSK19	**.68**	**.73**	.29	**.84**	.01	**.45**		**.83**		**.69**	**.90**		**.75**	.03	**.70**	**.83**		**.80**	.02
SCSK23	**.72**	**.66**	.38	**.66**	.09	**.30**		**.76**		**.78**	**.77**		**.63**	.13	**.74**	**.76**		**.67**	.13
SCSK26	**.73**	**.75**	.29	**.80**	−.04	**.35**		**.78**		**.66**	**.77**		**.77**	.00	**.65**	**.78**		**.81**	−.01
SCCH3	**.46**	**.54**	.14	**.56**	.01	**.06**		**.55**		**.69**	**.52**		**.52**	−.02	**.65**	**.51**		**.56**	−.03
SCCH7	**.36**	**.51**	.03	**.55**	−.02	**.02**		**.48**		**.62**	**.49**		**.50**	−.09	**.61**	**.43**		**.53**	−.10
SCCH10	**.44**	**.57**	.10	**.62**	.02	**.07**		**.56**		**.55**	**.56**		**.55**	−.04	**.54**	**.52**		**.58**	−.05
SCCH15	**.58**	**.66**	.19	**.68**	.00	**.08**		**.67**		**.67**	**.67**		**.65**	−.01	**.64**	**.64**		**.68**	−.02
SCMI9	**.50**	**.56**	.17	**.58**	−.01	**.09**		**.56**		**.75**	**.54**		**.54**	.00	**.69**	**.54**		**.57**	.00
SCMI14	**.59**	**.67**	.21	**.68**	.01	**−.02**		**.69**		**.72**	**.66**		**.65**	.01	**.67**	**.66**		**.68**	.01
SCMI17	**.61**	**.66**	.24	**.68**	.02	**.11**		**.68**		**.56**	**.66**		**.64**	.03	**.54**	**.66**		**.67**	.03
SCMI22	**.55**	**.66**	.17	**.75**	−.08	**.26**		**.66**		**.56**	**.69**		**.68**	−.06	**.56**	**.66**		**.73**	−.07
SCSJ1	**.67**	.21	**.77**	.00	**.79**		**−.16**		**.80**	**.43**		**.84**	.05	**.78**	**.60**		**.77**	−.05	.82
SCSJ8	**.66**	.19	**.69**	.04	**.74**		**.15**		**.72**	**.52**		**.84**	.03	**.70**	**.67**		**.74**	−.02	**.74**
SCSJ11	**.70**	.28	**.66**	.07	**.64**		**.09**		**.73**	**.53**		**.81**	.08	**.62**	**.70**		**.72**	.08	**.66**
SCSJ16	**.75**	.25	**.78**	.01	**.79**		**.06**		**.82**	**.55**		**.91**	.00	**.77**	**.75**		**.80**	.00	**.81**
SCSJ21	**.67**	.25	**.62**	.01	**.65**		**.34**		**.67**	**.52**		**.86**	.07	**.60**	**.75**		**.74**	.07	**.64**
SCIS4	**.66**	.22	**.70**	.01	**.75**		**.10**		**.73**	**.31**		**.87**	.01	**.70**	**.44**		**.72**	−.02	**.75**
SCIS13	**.64**	.23	**.62**	.03	**.66**		**.32**		**.67**	**.21**		**.81**	.04	**.62**	**.34**		**.67**	.04	**.65**
SCIS18	**.57**	.17	**.57**	.02	**.62**		**.32**		**.60**	**.29**		**.69**	.00	**.58**	**.43**		**.60**	−.01	**.61**
SCIS25	**.67**	.22	**.68**	.00	**.73**		**.18**		**.71**	**.40**		**.83**	.00	**.68**	**.58**		**.72**	.00	**.72**
SCOI2	**.69**	.22	**.78**	.00	**.84**		**−.07**		**.81**	**.34**		**.92**	.06	**.81**	**.48**		**.78**	−.06	**.85**

(*table continues*)

Table 5 (*continued*)

Self Compassion Scale (SCS) items	ESEM (Categorical)									Bayes (Continuous)					Bayes (Categorical)				
	Orthogonal			Correlated															
	6 + 1	6 + 2ESEM		6 + 2ESEM		6 + 2CFA		SET ESEM		6 + 1	6 + 2		6 + 2		6 + 1	6 + 2		6 + 2	
	GLB	CS	RUS	CS	RUS	CS	RUS	CS	RUS	GLB	CS	RUS	CS	RUS	GLB	CS	RUS	CS	RUS
SCOI6	**.68**	.21	**.73**	.02	**.78**		**.03**		**.76**	**.41**		**.87**	.03	**.74**	**.58**		**.75**	−.03	**.78**
SCOI20	**.59**	.20	**.58**	.00	**.60**		**.19**		**.61**	**.43**		**.65**	.00	**.58**	**.61**		**.59**	.01	**.61**
SCOI24	**.60**	.24	**.56**	.05	**.57**		**.21**		**.60**	**.38**		**.68**	.06	**.55**	**.56**		**.60**	.05	**.58**
M Target	**.62**	**.64**	**.67**	**.69**	**.70**	**.20**	**.14**	**.67**	**.71**	**.54**	**.69**	**.81**	**.64**	**.67**	**.61**	**.66**	**.71**	**.68**	**.71**

Note. Bold coefficients refer to factor loadings for items designed to reflect each factor (target loadings; nontarget loadings are unbolded). Coefficients shaded in gray are not statistically significant (or significantly negative). GLB = a single global factor; CS = compassionate self (based on positive SCS factors); RUS = uncompassionate self (based on negative SCS factors, but reverse scored); *M* Target = mean of target loadings. Orthogonal the Correlated refer to the rotation in the ESEMs.

To paraphrase a famous Mark Twain warning relevant to all data detectives, when the clock strikes 13, ignore the first 12.

Particularly for self-compassion research based on the SCS, our results clearly show we should not use 6ESEM + 2CFA. This is critical as interpretations based on this model are at the heart of recent debates about the structure of self-compassion that have important implications for theory, policy, and clinical practice. However, we leave as an open question as to whether this model is ever an appropriate approximation of the 6ESEM + 2ESEM model. Minimally researchers should scrutinize interpretations based on 6ESEM + 2CFA. However, we suggest that applied researchers use the 6Bayes + 2GlbBF model that can fit an appropriate bifactor model—one with six specific factors (and cross-loadings if applicable) and two global correlated bifactors. However, we also note that using CFA allows researchers to fit this model without cross-loadings. More complicated models are not always better if the results do not make sense and should always be juxtaposed with more parsimonious models.

Substantive Issues—One or Two Global Factors

Neff et al. (2019) and Rakhimov et al. (2023) both argue for a 6 + 1 model rather than a 6 + 2 model. However, the contention is based on a dubious interpretation of an inappropriate model. Nevertheless, our reanalysis shows that there is not much difference between the two models (6 + 1 and 6 + 2) when they are appropriately defined with Bayes. The 6 + 2 is stronger, but there is not much difference.

Neff et al. (2019) and Rakhimov et al. (2023) imply that the global factor in the 6 + 1 model is like a total score, and the six specific factors are like scale scores representing the six factors. However, this is not an appropriate representation of the 6 + 1 factors in a bifactor model. The one global score is like a total score AFTER controlling for variance explained by the six first-order factors. This is different from a global score without controlling for the six specific factors (i.e., a 1 + 0 model). Indeed, a model with just one global factor did not fit the data very well and was rejected by Neff et al. and Rakhimov et al. Similarly, the six specific factors in the 6 + 1 model are not the same as the six specific factors in a 6 + 0 model.

Bifactor models are useful to (a) separate item response variance into general versus specific group factor sources, (b) determine the degree that item responses conform to a unidimensional versus multidimensional structure, and (c) assess the utility of subscale scores after controlling variance due to the general factor (Reise, 2012). However, an ongoing problem with all the bifactor models (6 + 1 and 6 + 2) is how to interpret the results and what "scores" should be used to represent the SCS responses. The applied user and practitioner want a parsimonious, simple score (or set of scores) that is easily interpreted and diagnostic in practice. Maybe, based on the bifactor results, a reasonable compromise would be to present one or two global scores (or factor scores) based on a one-factor or a two-factor model and scale scores (or factor scores) based on a six-factor measurement model. These scores would be easily interpretable but are not scores generated from the bifactor model. It would be possible to generate factor scores from the bifactor model. However, these would be difficult for users (and even psychometricians) to interpret and, in many cases, would not be sufficiently reliable to warrant interpretation.

Neff (2022; also see Neff et al., 2019) argued for using a single global score in that the CS and RUS scales change in tandem. In support of this claim, Neff cited Ferrari et al.'s (2019) meta-analysis as evidence that all six subscales change simultaneously in response to training and interventions. We note, however, that most of the studies in the intervention were generic interventions targeting self-compassion broadly, not a specific component of self-compassion. If researchers want to evaluate SCS's discriminant validity with interventions, they need to consider multiple interventions designed to target specific factors. There would be support for the discriminant validity of factors if the change in the targeted factor is consistently larger than changes in the nontargeted factors, and this pattern of results is consistent over multiple interventions that target different specific factors (e.g., Marsh & Roche, 1993, 1997).

Adopting a different approach to the relation between CS and RUS, Mantzios et al. (2020) created opposite versions of the CS items. For example, they rephrased self-kindness to measure self-unkindness. These novels, rephrased items had strong positive correlations with the original RUS items, suggesting the CS and RUS subscales are largely equivalent. We note, however, that our results (Table 4) did not suggest that matching CS and RUS components correlated more highly than nonmatching CS and RUS components. Mantzios et al. also examined the effects of an intervention targeting an increase in self-compassion, with an intervention targeting a decrease in self-criticism. Differences across key psychological outcomes, including total SCS score, were negligible. Significant differences emerged, however, when examining the six subscales, with MI significantly improving only in the self-compassionate condition, and self-judgment only in the self-critical condition. These trends support the discriminant validity of the six SCS components that would be lost if only a total score is considered.

We also worry that the single global score seems to deny the possibility that a person can be high on both CS and RUS, or low on both CS and RUS. SCS research routinely reverse-scores RUS (reverse-scored uncompassionate self-responding). Hence, a high (reverse-scored) RUS score represents uncompassionate self-responding. Therefore, combining CS and (reverse-scored) RUS into a single global score reflects a bipolar perspective of self-compassion. Thus, according to this bipolar representation, a person very high on compassionate self-responding must necessarily be very low on RUS. Some studies have used profile analysis of the SCS six subscores to identify groups that tend to respond to all items of the SCS uniformly. For example, Ferrari et al. (2022) found adolescent males either weakly endorsed all six subscales; or moderately endorsed all six subscales, and 16% of female adolescents strongly endorsed both SC and RUS. In a Chinese university sample, Wu et al. (2020) found two of their four profiles in which students either moderately endorsed both SC and RUS ("nondialectical high self-compassion") or weakly endorsed both ("nondialectical low self-compassion"). In contrast, the two-global-factor interpretation would resolve this issue. Maybe this occurrence is unusual, particularly in the extreme (e.g., Phillips, 2021; Ullrich-French & Cox, 2020). However, it would seem to be important from a clinical perspective to develop an accurate case formulation of the patient's presenting problem, and to guide evidence-based treatment planning for psychological intervention.

Theoretical Issues—The Six First-Order ("Specific") Factors

For us, the critical issue is the theoretical basis for the six first-order ("specific") SCS factors. Regarding diagnostic value, the global scores (one or two) might represent a "red flag" in the sense of giving a global warning. However, the real diagnostic value to a clinician should focus on the specific first-order factors—particularly if one stands out as different from the others. If the first-order factors are sufficiently distinct, then it is likely that interventions that focus on one of the specific factors will have more effect on that factor than on other factors. Indeed, this is a strong test of the discriminant validity of the different factors. Hence, the instrument must contain the most relevant specific factors. This is more important than the one- versus two-global factor debate.

We are not arguing that the six factors in the SCS are inappropriate. Nor do we contend that additional factors should be included instead of, or in addition to, the six SCS factors. Instead, we argue that it would be more fruitful to critically evaluate the theoretical underpinning of the self-compassion construct that should be the basis for selecting first-order factors. Thus, for example, Gilbert and colleagues (Gilbert, 2010, 2014; Gilbert & Simos, 2022) take a very different approach to the underlying constructs (e.g., noncondemning/nonjudging; sympathy, empathy) that make up the self-compassion construct.

Neff et al. (2019; Neff, 2023) argue for a bipolar representation of self-compassion in which CS components define the positive pole, and RUS components define the negative pole. In their theoretical model, each CS factor has an implicit pairing with a matching RUS. For example, in advice for clinical practice, Germer and Neff (2013; also see Neff, 2023) noted that For example, in advice for clinical practice, Germer and Neff (2013; also see Neff, 2023) noted that "Self-compassion is conceptualized as containing 3 core components: self-kindness versus self-judgment, common humanity versus isolation, and mindfulness versus overidentification, when relating to painful experiences (p. 856). Similarly, Neff (2023) noted that she operationalized self-compassion as a multifaceted construct comprising three broad domains that are overlapping but conceptually distinct. However, it is unclear whether this is a critical theoretical feature of self-compassion theory in which the three pairs of factors define three bipolar components of self-compassion. Indeed, it is surprising that models with three bipolar factors are not featured in SCS research. Nevertheless, the pattern of correlations among the six SCS factors (Table 4; but also see Neff et al., 2018, 2019) offered little support for this theoretical pairing of positive and negative factors. For this reason—and because the focus was on the 6ESEM + 2CFA model—we did not pursue models positing three bipolar factors. However, these preliminary results seem inconsistent with Neff et al.'s emphasis on bipolar factors.

Neff et al. (2019) have not given sufficient attention to the overarching problem of whether the distinction between negatively and positively oriented constructs reflects a substantively important distinction, a methodological artifact associated with item coding, or a combination of the two. This is a very general problem in psychological assessment with no easy solution. Indeed, Murphy et al. (2023; also see Böckenholt, 2019) recently referred to this issue as the "item coding direction: The elephant in the room" (p. 4). The insertion of negatively coded items is widely recommended to disrupt response acquiescence and careless responding (e.g., Nunnally, 1994). However, this practice invariably results in poorly defined factors when positively and negatively worded items are posited to load on the same factor. In other areas, researchers have taken different perspectives on this issue. Thus, for example, based on the heated debate over the structure of the Rosenberg self-esteem instrument, Marsh et al. (2010) argued for a unidimensional (bipolar) representation that included method effects for negatively worded items. However, Pekrun et al. (2023) argued that valence is a substantively important component in their multidimensional structure of emotions. Marsh et al. (2013) used a construct validity approach to support the separation of passion into harmonious- and obsessive-passion factors. Dweck and Leggett's (1988) implicit theory of intelligence posited a bipolar construct with fixed and growth-mindset poles, but empirical support is mixed. Indeed, some implicit theory researchers finesse the issue by only presenting items from the fixed end of the bipolar continuum (but still referring to the construct as growth mindset rather than fixed mindset). Ryan and Deci (2017) argue for the substantive importance of distinguishing between need satisfaction and need frustration (but see Murphy et al., 2023; Tóth-Király et al., 2018). Indeed, the positive psychology movement contends that positive mental health is not just the absence of mental illness (Marsh, Huppert, et al., 2020).

A discussion of the item-coding effect is beyond the scope of the present investigation, but it is relevant to interpreting SCS responses. The SCS scales completely confound the effects of item-coding, in that the positive components are all defined by positively coded items, and negative components are all defined by reverse-coded items. Hence, to the extent that there is an item-coding method effect, it is likely to make positive and negative components of compassion artificially more distinct. From this perspective, it might be reasonable to argue that the .6 correlation between positive and negative components of self-compassion is consistent with a single substantive global component that is distorted by an item-coding method effect. This would be consistent with Neff et al.'s (2019) support for a single global component. However, a stronger test would require an unconfounding direction of item-wording within each SCS. Nevertheless, this strategy would probably undermine the six-factor model's clean factor structure and good fit. There is no easy solution to this issue.

Our purpose here is not to resolve these critical theoretical issues but to highlight ones needing attention. A focus on critical theoretical issues seems to have been side-tracked by statistical concerns and the debate about one- versus two-global-factor models. Psychometric support is vital but not a replacement for strong theory.

A Rapprochement: Does Neff Reject Models With Two (CS and RUS) Global Constructs?

Implicit in much of the debate on SCS's global structure is the assumption that Neff and colleagues reject all models with two (CS and RUS) global constructs. However, this is not accurate. Rather, Neff et al. (2018, 2019) rejected the 6ESEM + 2CFA model. Indeed, we also reject this model not because it posits two global factors but because it is flawed. However, rejecting that particular model does not mean rejecting all structures with two global factors. Thus, the 6CFA + 2GlbBF and particularly the 6Bayes + 2GlbBF models fit the data well and provide internally consistent parameter estimates. Furthermore, rejecting all models with two global factors is inconsistent with much of what Neff has actually written about SCS's structure. For example, Neff et al.'s (2018) results support separating the two global (CS and RUS) factors; they offer substantive rationales for why CS and RUS are differentially related to different criterion variables. Neff (2023) noted the importance of establishing norms and clinical cut-off values for the SCS. However, she emphasized that the question of whether establishing clinical norms is best done with a total score, the six subscale scores, or even two separate CS and RUS scores is yet to be answered. Finally, Neff argued that how the SCS is used should depend on the purposes of the researcher (Neff et al., 2021).

Indeed, commenting on an earlier draft of this study (January 08, 2023, personal communication), Neff wrote:

> I don't have a problem with researchers using either a single score or two separate CS and RUS scores, I mainly have a problem with people who argue that one should never use a total score. First, your analyses support a position that one or two scores are pretty much equally valid, and would not support the view that one should never use a total score. I would argue that use of a total score, two CS and RUS scores, or six subscales scores should be determined by the purposes of the researcher.

We agree.

In summary, we see reasonable agreement between Neff (2016; Neff et al., 2018, 2019) and our study. Both reject the 6ESEM + 2CFA model for substantially overlapping reasons. Neff et al. (2019) argue that the model provides weak evidence supporting two global factors. We agree but contend that the problem is an inappropriate model rather than a structure with two global factors. Although our results favor models with CS and RUS factors over models with a total SCS score, the difference in fit between Bayes representations (6Bayes + 2GlbBF vs. 6Bayes + 1GlbBF) is small. Hence, we agree with Neff that using a total SCS score or separate CS and RUS scores might be appropriate in some circumstances.

Conclusions and Implications

Self-compassion has a long history in Eastern Philosophies but has experienced a surge in clinical psychology due to its association with many mental health outcomes. Self-compassion research is driven mainly by Neff's (2003a, 2003b) SCS, but there is much debate on SCS's global structure (one- vs. two-global factors) and what SCS scores should be used. Neff et al. (2019) compared ESEM models with one and two global factors, arguing for a one-global factor solution. However, ESEM did not allow them to test the appropriate bifactor model with two global factors, and their intuitively reasonable approximation of the appropriate model resulted in internally inconsistent, illogical results.

In our reanalysis of Neff et al.'s (2019) data, we used evolving Bayes estimation to fit more appropriate bifactor models. The Bayes model with two global factors performs slightly better than the one-global factor model but the differences are not substantial. We discuss the critical implications for theory, scoring, and application of the SCS inappropriately based on the discredited ESEM model used by Neff et al. For application in applied clinical and research settings, we endorse using all the scores representing the six SCS factors, total SCS, compassionate self-responding (CS), and RUS rather than relying solely on one global factor in IS.

More broadly, developing and applying clinician instruments requires considering their dimensionality and factor structures. By understanding a measure's factor structure, clinicians and researchers can accurately diagnose and treat individuals based on their unique factor profiles. For example, when a client is high on compassion and low on uncompassionate behavior, self-compassion interventions might be less useful as they already show an "ideal" profile. Alternatively, if someone has high levels of both uncompassionate and compassionate responses, interventions may focus on reducing uncompassionate responses rather than increasing the already high frequency of compassionate behavior. Identifying factor structure allows clinical researchers to hypothesize and test them in future research.

A variety of clinically relevant instruments distinguish between positively and negatively valenced items. Theoretically, such items may link to behavioral activation versus behavioral inhibition biological systems (Carver & White, 1994), behavior that promote positive outcomes versus behaviors that creates problems (Ciarrochi et al., 2022), positive and negative affectivity (P. Costa & McCrae, 1992; Tellegen et al., 1999), and need satisfaction and frustration (Chen et al., 2015). It is also possible to empirically distinguish between positive interventions, such as behavioral activation, and more negatively focused interventions, such as reappraising negative thoughts (Ciarrochi et al., 2021; Dimidjian et al., 2006).

Ultimately there is a need for substantive-methodological synergy (Marsh & Hau, 2007) that combines theory, measurement, and statistical analysis in a helpful way for research, intervention, policy, and practice. In their manifesto on substantive-methodological synergy, Marsh and Hau (2007) argued that applied researchers applying new and evolving methodologies should adopt the role of data detective. They should thoroughly evaluate the appropriateness of new methodological approaches and interpretations using a construct validity approach based on theory. Our study demonstrates this approach to assess the application of 6ESEM + 2CFA in self-compassion research and why this is important. However, the study also is an exemplar for many clinical measurement studies facing similar issues. Issues raised here (dimensionality, factor structure, first-order and higher order models, positive vs. negatively oriented constructs, item-wording effects, and alternative estimation procedures) have wide applicability to clinical measurement (see supplemental material Section 1 where we annotate 20 clinical instruments that might benefit from approaches used here).

References

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, *16*(3), 397–438. https://doi.org/10.1080/10705510903008204

Asparouhov, T., & Muthén, B. (2021). Advances in Bayesian model fit evaluation for structural equation models. *Structural Equation Modeling*, *28*(1), 1–14. https://doi.org/10.1080/10705511.2020.1764360

Bandalos, D. L. (2014). Relative performance of categorical diagonally weighted least squares and robust maximum likelihood estimation. *Structural Equation Modeling*, *21*(1), 102–116. https://doi.org/10.1080/10705511.2014.859510

Böckenholt, U. (2019). Contextual responses to affirmative and/or reversed-worded items. *Psychometrika*, *84*(4), 986–999. https://doi.org/10.1007/s11336-019-09680-7

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. https://doi.org/10.1037/cou0000211

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology*, *67*(2), 319–333. https://doi.org/10.1037/0022-3514.67.2.319

Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E. L., Van der Kaap-Deeder, J., Duriez, B., Lens, W., Matos, L., Mouratidis, A., Ryan, R. M., Sheldon, K. M., Soenens, B., Van Petegem, S., & Verstuyf, J. (2015). Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion*, *39*(2), 216–236. https://doi.org/10.1007/s11031-014-9450-1

Ciarrochi, J., Hayes, S. C., Oades, L. G., & Hofmann, S. G. (2021). Toward a Unified Framework for Positive Psychology Interventions: Evidence-Based Processes of Change in Coaching, Prevention, and Training. *Frontiers in Psychology*, *12*, Article 809362. https://doi.org/10.3389/fpsyg.2021.809362

Ciarrochi, J., Sahdra, B., Hofmann, S. G., & Hayes, S. C. (2022). Developing an item pool to assess processes of change in psychological interventions: The Process-Based Assessment Tool (PBAT). *Journal of Contextual Behavioral Science*, *23*, 200–213. https://doi.org/10.1016/j.jcbs.2022.02.001

Cleare, S., Gumley, A., Cleare, C. J., & O'Connor, R. C. (2018). An investigation of the factor structure of the self-compassion scale. *Mindfulness*, *9*(2), 618–628. https://doi.org/10.1007/s12671-017-0803-1

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. https://doi.org/10.1002/cpp.1974

Costa, P., & McCrae, R. (1992). *The NEO Personality Inventory Manual*. Psychological Assessment Resources.

Dimidjian, S., Hollon, S. D., Dobson, K. S., Schmaling, K. B., Kohlenberg, R. J., Addis, M. E., Gallop, R., McGlinchey, J. B., Markley, D. K., Gollan, J. K., Atkins, D. C., Dunner, D. L., & Jacobson, N. S. (2006). Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *Journal of Consulting and Clinical Psychology*, *74*(4), 658–670. https://doi.org/10.1037/0022-006X.74.4.658

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*(2), 256–273. https://doi.org/10.1037/0033-295X.95.2.256

Fan, Y.-C., Hsiao, F.-H., & Hsieh, C.-C. (2023). The effectiveness of compassion-based interventions among cancer patients: A systematic review and meta-analysis. *Palliative & Supportive Care*, *21*(3), 534–546. https://doi.org/10.1017/S1478951522001316

Ferrari, M., Ciarrochi, J., Yap, K., Sahdra, B., & Hayes, S. C. (2022). Embracing the complexity of our inner worlds: Understanding the dynamics of self-compassion and self-criticism. *Mindfulness*, *13*(7), 1–10. https://doi.org/10.1007/s12671-022-01897-5

Ferrari, M., Hunt, C., Harrysunker, A., Abbott, M. J., Beath, A. P., & Einstein, D. A. (2019). Self-compassion interventions and psychosocial outcomes: A meta-analysis of RCTs. *Mindfulness*, *10*(8), 1455–1473. https://doi.org/10.1007/s12671-019-01134-6

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis*. Chapman & Hall/CRC.

Germer, C. K., & Neff, K. D. (2013). Self-compassion in clinical practice. *Journal of Clinical Psychology*, *69*(8), 856–867. https://doi.org/10.1002/jclp.22021

Gilbert, P. (2010). *Compassion focused therapy: Distinctive features*. Routledge/Taylor & Francis Group.

Gilbert, P. (2014). The origins and nature of compassion focused therapy. *British Journal of Clinical Psychology*, *53*(1), 6–41. https://doi.org/10.1111/bjc.12043

Gilbert, P., McEwan, K., Matos, M., & Rivis, A. (2011). Fears of compassion: Development of three self-report measures. *Psychology and Psychotherapy: Theory, Research and Practice*, *84*(3), 239–255. https://doi.org/10.1348/147608310X526511

Gilbert, P., & Simos, G. (Eds.). (2022). *Compassion focused therapy: Clinical practice and applications*. Routledge. https://doi.org/10.4324/9781003035879

Guo, J., Marsh, H. W., Parker, P. D., Dicke, T., Ludtke, O., & Diallo, T. M. O. (2019). A systematic evaluation and € comparison between exploratory structural equation modeling and Bayesian structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(4), 529–556. https://doi.org/10.1080/10705511.2018.1554999

Helminen, E. C., Ducar, D. M., Scheer, J. R., Parke, K. L., Morton, M. L., & Felver, J. C. (2023). Self-compassion, minority stress, and mental health in sexual and gender minority populations: A meta-analysis and systematic review. *Clinical Psychology: Science and Practice*, *30*(1), 26–39. https://doi.org/10.1037/cps0000104

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

Inwood, E., & Ferrari, M. (2018). Mechanisms of change in the relationship between self-compassion, emotion regulation, and mental health: A systematic review. *Applied Psychology. Health and Well-Being*, *10*(2), 215–235. https://doi.org/10.1111/aphw.12127

Kariyawasam, L., Ononaiye, M., Irons, C., & Kirby, S. E. (2023). Compassion-based interventions in Asian communities: A meta-analysis of randomised controlled trials. *Psychology and Psychotherapy: Theory, Research and Practice*, *96*(1), 148–171. https://doi.org/10.1111/papt.12431

Kirby, J. N., Tellegen, C. L., & Steindl, S. R. (2017). A meta-analysis of compassion-based interventions: Current state of knowledge and future directions. *Behavior Therapy*, *48*(6), 778–792. https://doi.org/10.1016/j.beth.2017.06.003

Kumlander, S., Lahtinen, O., Turunen, T., & Salmivalli, C. (2018). Two is more valid than one, but is six even better? The factor structure of the Self-Compassion Scale (SCS). *PLOS ONE*, *13*(12), Article e0207706. https://doi.org/10.1371/journal.pone.0207706

Larsen, R. J., & Diener, E. (1987). Affect intensity as an individual difference characteristic: A review. *Journal of Research in Personality*, *21*(1), 1–39. https://doi.org/10.1016/0092-6566(87)90023-7

MacBeth, A., & Gumley, A. (2012). Exploring compassion: A meta-analysis of the association between self-compassion and psychopathology. *Clinical Psychology Review*, *32*(6), 545–552. https://doi.org/10.1016/j.cpr.2012.06.003

Mantzios, M., Koneva, A., & Egan, H. (2020). When "negativity" becomes obstructive: A novel exploration of the two-factor model of the self-compassion scale and a comparison of self-compassion and self-criticism interventions. *Current Issues in Personality Psychology*, *8*(4), 289–300. https://doi.org/10.5114/cipp.2020.100791

Marsh, H. W., Byrne, B. M., & Yeung, A. S. (1999). Causal ordering of academic self-concept and achievement: Reanalysis of a pioneering study and. *Educational Psychologist*, *34*(3), 155–167. https://doi.org/10.1207/s15326985ep3403_2

Marsh, H. W., Guo, J., Dicke, T., Parker, P. D., & Craven, R. G. (2020). Confirmatory factor analysis (CFA), exploratory structural equation modeling (ESEM), and set-ESEM: Optimal balance between goodness of fit and parsimony. *Multivariate Behavioral Research*, *55*(1), 102–119. https://doi.org/10.1080/00273171.2019.1602503

Marsh, H. W., & Hau, K.-T. (2007). Applications of latent-variable models in educational psychology: The need for methodological-substantive synergies. *Contemporary Educational Psychology*, *32*(1), 151–170. https://doi.org/10.1016/j.cedpsych.2006.10.008

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal*, *11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2

Marsh, H. W., Huppert, F. A., Donald, J. N., Horwood, M. S., & Sahdra, B. K. (2020). The well-being profile (WB-Pro): Creating a theoretically based multidimensional measure of well-being to advance theory, research, policy, and practice. *Psychological Assessment*, *32*(3), 294–313. https://doi.org/10.1037/pas0000787

Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology*, *10*(1), 85–110. https://doi.org/10.1146/annurev-clinpsy-032813-153700

Marsh, H. W., & Roche, L. (1993). The use of students' evaluations and an individually structured intervention to enhance university teaching effectiveness. *American Educational Research Journal*, *30*(1), 217–251. https://doi.org/10.3102/00028312030001217

Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist*, *52*(11), 1187–1197. https://doi.org/10.1037/0003-066X.52.11.1187

Marsh, H. W., Scalas, L. F., & Nagengast, B. (2010). Longitudinal tests of competing factor structures for the Rosenberg Self-Esteem Scale: Traits, ephemeral artifacts, and stable response styles. *Psychological Assessment*, *22*(2), 366–381. https://doi.org/10.1037/a0019225

Marsh, H. W., Vallerand, R. J., Lafrenière, M.-A. K., Parker, P., Morin, A. J. S., Carbonneau, N., Jowett, S., Bureau, J. S., Fernet, C., Guay, F., Abduljabbar, A. S., & Paquet, Y. (2013). Passion: Does one scale fit all? Construct validity of two-factor passion scale and psychometric invariance over different activities and languages. *Psychological Assessment*, *25*(3), 796–809. https://doi.org/10.1037/a0032573

Morin, A. J. S., Arens, A. K., & Marsh, H. W. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling*, *23*(1), 116–139. https://doi.org/10.1080/10705511.2014.961800

Morin, A. J. S., Boudrias, J.-S., Marsh, H. W., Madore, I., & Desrumaux, P. (2016). Further reflections on disentangling shape and level effects in person-centered analyses: An illustration exploring the dimensionality of psychological health. *Structural Equation Modeling*, *23*(3), 438–454. https://doi.org/10.1080/10705511.2015.1116077

Morin, A. J. S., Myers, N. D., & Lee, S. (2020). Modern factor analytic techniques. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology*, (pp. 1044–1073). Wiley. https://doi.org/10.1002/9781119568124.ch51

Muris, P. (2016). A protective factor against mental health problems in youths? A critical note on the assessment of self-compassion. *Journal of Child and Family Studies*, *25*(5), 1461–1465. https://doi.org/10.1007/s10826-015-0315-3

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. https://doi.org/10.1007/s12671-020-01363-0

Muris, P., Otgaar, H., López, A., Kurtic, I., & van de Laar, I. (2021). The (non) protective role of self-compassion in internalizing symptoms: Two empirical studies in adolescents demonstrating unwanted effects of using the self-compassion scale total score. *Mindfulness*, *12*, 240–252. https://doi.org/10.1007/s12671-020-01514-3

Muris, P., Otgaar, H., & Pfattheicher, S. (2019). Stripping the forest from the rotten trees: Compassionate self-responding is a way of coping, but reduced uncompassionate self-responding mainly reflects psychopathology. *Mindfulness*, *10*(1), 196–199. https://doi.org/10.1007/s12671-018-1030-0

Muris, P., & Petrocchi, N. (2017). Protection or vulnerability? A meta-analysis of the relations between the positive and negative components of self-compassion and psychopathology. *Clinical Psychology & Psychotherapy*, *24*(2), 373–383. https://doi.org/10.1002/cpp.2005

Murphy, B. A., Watts, A. L., Baker, Z. G., Don, B. P., Jolink, T. A., & Algoe, S. B. (2023). The Basic Psychological Need Satisfaction and Frustration Scales probably do not validly measure need frustration. *Psychological Assessment*, *35*(2), 127–139. https://doi.org/10.1037/pas0001193

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, *17*(3), 313–335. https://doi.org/10.1037/a0026802

Muthén, L. K., & Muthén, B. O. (1998–2018). *Mplus user's guide* (8th ed.).

Neff, K. D. (2003a). The development and validation of a scale to measure self-compassion. *Self and Identity*, *2*(3), 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*(2), 85–101. https://doi.org/10.1080/15298860309032

Neff, K. D. (2016). Does self-compassion entail reduced self-judgment, isolation, and over-identification? A response to Muris, Otgaar, and Petrocchi (2016). *Mindfulness*, *7*(3), 791–797. https://doi.org/10.1007/s12671-016-0531-y

Neff, K. D. (2022). The differential effects fallacy in the study of self-compassion: Misunderstanding the nature of bipolar continuums. *Mindfulness*, *13*(3), 572–576. https://doi.org/10.1007/s12671-022-01832-8

Neff, K. D. (2023). Self-compassion: Theory, method, research, and intervention. *Annual Review of Psychology*, *74*(1), 193–218. https://doi.org/10.1146/annurev-psych-032420-031047

Neff, K. D., Long, P., Knox, M., Davidson, O., Kuchar, A., Costigan, A., Williamson, Z., Rohleder, N., Tóth-Király, I., & Breines, J. (2018). The forest and the trees: Examining the association of self-compassion and its positive and negative components with psychological functioning. Self and Identity. *Self and Identity*, *17*(6), 627–645. https://doi.org/10.1080/15298868.2018.1436587

Neff, K. D., Tóth-Király, I., Knox, M. C., Kuchar, A., & Davidson, O. (2021). The development and validation of the State Self-Compassion Scale (Long-and Short Form). *Mindfulness*, *12*(1), 121–140. https://doi.org/10.1007/s12671-020-01505-4

Neff, K. D., Tóth-Király, I., Yarnell, L. M., Arimitsu, K., Castilho, P., Ghorbani, N., Guo, H. X., Hirsch, J. K., Hupfeld, J., Hutz, C. S., Kotsou, I., Lee, W. K., Montero-Marin, J., Sirois, F. M., de Souza, L. K., Svendsen, J. L., Wilkinson, R. B., & 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. https://doi.org/10.1037/pas0000629

Nunnally, J. C. (1994). *Psychometric theory 3E*. Tata McGraw-Hill Education.

Pekrun, R., Marsh, H. W., Elliot, A. J., Stockinger, K., Perry, R. P., Vogl, E., Goetz, T., van Tilburg, W. A. P., Lüdtke, O., & Vispoel, W. P. (2023). A three-dimensional taxonomy of achievement emotions. *Journal of Personality and Social Psychology*, *124*(1), 145–178. https://doi.org/10.1037/pspp0000448

Phillips, W. J. (2021). Self-compassion mindsets: The components of the self-compassion scale operate as a balanced system within individuals. *Current Psychology*, *40*(10), 5040–5053. https://doi.org/10.1007/s12144-019-00452-1

Phillips, W. J., & Hine, D. W. (2021). Self-compassion, physical health, and health behaviour: A meta-analysis. *Health Psychology Review*, *15*(1), 113–139.

Rakhimov, A., Realo, A., & Tang, N. K. Y. (2023). The self-compassion scale: Validation and psychometric properties within the exploratory structural equation modeling framework. *Journal of Personality Assessment*, *105*(3), 422–435. https://doi.org/10.1080/00223891.2022.2093731

Reise, S. P. (2012). Invited paper: The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*(5), 667–696. https://doi.org/10.1080/00273171.2012.715555

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Press. https://doi.org/10.1521/978.14625/28806

Sass, D. A., Schmitt, T. A., & Marsh, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators. *Structural Equation Modeling*, *21*(2), 167–180. https://doi.org/10.1080/10705511.2014.882658

Simpson, R., Posa, S., Bruno, T., Simpson, S., Wasilewski, M. B., Robinson, L. R., Munce, S., Bayley, M., & Feinstein, A. (2023). Conceptualization, use, and outcomes associated with compassion in the care of people with multiple sclerosis: A scoping review. *Journal of Neurology*, *270*(3), 1300–1322. https://doi.org/10.1007/s00415-022-11497-x

Tellegen, A., Watson, D., & Clark, L. A. (1999). On the dimensional and hierarchical structure of affect. *Psychological Science*, *10*(4), 297–303. https://doi.org/10.1111/1467-9280.00157

Tóth-Király, I., Morin, A. J., Bőthe, B., Orosz, G., & Rigó, A. (2018). Investigating the multidimensionality of need fulfillment: A bifactor exploratory structural equation modeling representation. *Structural Equation Modeling*, *25*(2), 267–286. https://doi.org/10.1080/10705511.2017.1374867

Tóth-Király, I., & Neff, K. D. (2021). Is self-compassion universal? Support for the measurement invariance of the self-compassion scale across populations. *Assessment*, *28*(1), 169–185. https://doi.org/10.1177/1073191120926232

Turk, F., & Waller, G. (2020). Is self-compassion relevant to the pathology and treatment of eating and body image concerns? A systematic review and meta-analysis. *Clinical Psychology Review*, *79*, Article 101856. https://doi.org/10.1016/j.cpr.2020.101856

Ullrich-French, S., & Cox, A. E. (2020). The use of latent profiles to explore the multi-dimensionality of self-compassion. *Mindfulness*, *11*(6), 1483–1499. https://doi.org/10.1007/s12671-020-01365-y

Wakelin, K. E., Perman, G., & Simonds, L. M. (2022). Effectiveness of self-compassion-related interventions for reducing self-criticism: A systematic review and meta-analysis. *Clinical Psychology & Psychotherapy*, *29*(1), 1–25. https://doi.org/10.1002/cpp.2586

Wechsler, D. (2014). *Wechsler intelligence scale for children* (5th ed.). Pearson.

Wu, Q., Chen, C., Liang, Y., Zhou, N., Cao, H., Du, H., Lin, X., & Chi, P. (2020). Not only the forest and trees but also the ground they are rooted in: Identifying profiles of self-compassion from the perspective of dialecticism. *Mindfulness*, *11*(8), 1967–1977. https://doi.org/10.1007/s12671-020-01406-6

Received January 12, 2023
Revision received April 7, 2023
Accepted April 28, 2023 ■