

Psychological Assessment

Examining the Factor Structure of the Self-Compassion Scale in 20 Diverse Samples: Support for Use of a Total Score and Six Subscale Scores

Kristin D. Neff, István Tóth-Király, Lisa M. Yarnell, Kohki Arimitsu, Paula Castilho, Nima Ghorbani, Hailan Xiaoxia Guo, Jameson K. Hirsch, Jörg Hupfeld, Claudio S. Hutz, Ilios Kotsou, Woo Kyeong Lee, Jesus Montero-Marin, Fuschia M. Sirois, Luciana K. de Souza, Julie L. Svendsen, Ross B. Wilkinson, and Michail Mantzios

Online First Publication, August 20, 2018. <http://dx.doi.org/10.1037/pas0000629>

CITATION

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. (2018, August 20). 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*. Advance online publication. <http://dx.doi.org/10.1037/pas0000629>

Examining the Factor Structure of the Self-Compassion Scale in 20 Diverse Samples: Support for Use of a Total Score and Six Subscale Scores

Kristin D. Neff
University of Texas at Austin

István Tóth-Király
ELTE Eötvös Loránd University

Lisa M. Yarnell
American Institutes for Research, Washington, DC

Kohki Arimitsu
Kwansei Gakuin University

Paula Castilho
University of Coimbra

Nima Ghorbani
University of Tehran

Hailan Xiaoxia Guo
Beijing Hailan Peer Education & Consultation Co., Beijing,
China

Jameson K. Hirsch
East Tennessee State University

Jörg Hupfeld
University of Bern

Claudio S. Hutz
Universidade Federal do Rio Grande do Sul

Ilios Kotsou
Université Libre de Bruxelles

Woo Kyeong Lee
Seoul Cyber University

Jesus Montero-Marin
Primary Care Prevention and Health Promotion Research
Network (RedIAPP), Zaragoza, Spain

Fuschia M. Sirois
University of Sheffield

Luciana K. de Souza
Universidade Federal do Rio Grande do Sul

Julie L. Svendsen
University of Bergen

Ross B. Wilkinson
University of Newcastle

Michail Mantzios
Birmingham City University

Kristin D. Neff, Department of Educational Psychology, University of Texas at Austin; István Tóth-Király, Doctoral School of Psychology, Department of Personality and Health Psychology, ELTE Eötvös Loránd University; Lisa M. Yarnell, American Institutes for Research, Washington, DC; Kohki Arimitsu, Department of Psychological Sciences, Kwansei Gakuin University; Paula Castilho, Faculty of Psychology and Educational Sciences, University of Coimbra; Nima Ghorbani, Department of Psychology, University of Tehran; Hailan Xiaoxia Guo, Beijing Hailan Peer Education & Consultation Co., Beijing, China; Jameson K. Hirsch, Department of Psychology, East Tennessee State University; Jörg Hupfeld, Department of Psychology, University of Bern; Claudio S. Hutz, Post-Graduate Program in Psychology, Universidade Federal do Rio Grande do Sul; Ilios Kotsou, Department of Psychology, Université Libre de Brux-

elles; Woo Kyeong Lee, Department of Counseling Psychology, Seoul Cyber University; Jesus Montero-Marin, Primary Care Prevention and Health Promotion Research Network (RedIAPP), Zaragoza, Spain; Fuschia M. Sirois, Department of Psychology, University of Sheffield; Luciana K. de Souza, Post-Graduate Program in Psychology, Universidade Federal do Rio Grande do Sul; Julie L. Svendsen, Department of Biological and Medical Psychology, University of Bergen; Ross B. Wilkinson, School of Psychology, University of Newcastle; Michail Mantzios, Department of Psychology, Birmingham City University.

Kristin D. Neff conceived the article and wrote the first draft. István Tóth-Király conducted all statistical analyses. Lisa M. Yarnell coordinated data collection and contributed her statistical expertise. Kristin D. Neff, István Tóth-Király, and Lisa M. Yarnell did the bulk of the writing. Michail Mantzios, coordinated putting together the translation information Table in the

continued

This study examined the factor structure of the Self-Compassion Scale (SCS) using secondary data drawn from 20 samples ($N = 11,685$)—7 English and 13 non-English—including 10 community, 6 student, 1 mixed community/student, 1 meditator, and 2 clinical samples. Self-compassion is theorized to represent a system with 6 constituent components: self-kindness, common humanity, mindfulness and reduced self-judgment, isolation and overidentification. There has been controversy as to whether a total score on the SCS or if separate scores representing compassionate versus uncompassionate self-responding should be used. The current study examined the factor structure of the SCS using confirmatory factor analyses (CFA) and Exploratory Structural Equation Modeling (ESEM) to examine 5 distinct models: 1-factor, 2-factor correlated, 6-factor correlated, single-bifactor (1 general self-compassion factor and 6 group factors), and 2-bifactor models (2 correlated general factors each with 3 group factors representing compassionate or uncompassionate self-responding). Results indicated that a 1- and 2-factor solution to the SCS had inadequate fit in every sample examined using both CFA and ESEM, whereas fit was excellent using ESEM for the 6-factor correlated, single-bifactor and correlated 2-bifactor models. However, factor loadings for the correlated 2-bifactor models indicated that 2 separate factors were not well specified. A general factor explained 95% of the reliable item variance in the single-bifactor model. Results support use of the SCS to examine 6 subscale scores (representing the constituent components of self-compassion) or a total score (representing overall self-compassion), but not separate scores representing compassionate and uncompassionate self-responding.

Public Significance Statement

This study examined the factor structure of the Self-Compassion Scale (SCS) in 20 diverse samples ($N = 11,685$), and excellent fit was found in every sample for an Exploratory Structural Equation Modeling (ESEM) single-bifactor model (with 95% of item variance explained by a general factor) and an ESEM 6-factor correlated model. Results support use of a total SCS score or 6 subscale scores, but not 2 separate scores representing compassionate and uncompassionate self-responding.

Keywords: Self-Compassion Scale, self-compassion factor structure, bifactor analyses, Exploratory Structural Equation Modeling (ESEM)

Supplemental materials: <http://dx.doi.org/10.1037/pas0000629.supp>

The construct of self-compassion was first operationally defined and introduced into the psychological literature a decade and a half ago (Neff, 2003b). Theoretically, self-compassion is comprised of six components that combine and mutually interact to create a self-compassionate frame of mind when faced with personal inadequacy or life difficulties: self-kindness versus self-judgment, a sense of common humanity versus isolation, and mindfulness versus overidentification. Self-kindness entails being gentle, supportive, and understanding toward oneself. Rather than harshly judging oneself for shortcomings, the self is offered warmth and acceptance. Common humanity involves recognizing the shared human experience, understanding that all humans fail, make mistakes, and lead imperfect lives. Rather than feeling isolated by one's imperfection—egocentrically feeling as if "I" am the only one who has failed or am suffering—one takes a broader and more connected perspective with regard to personal shortcomings and individual difficulties. Mindfulness involves being aware of one's present moment experience of suffering with clarity and balance, without running away with a dramatic storyline about negative

aspects of oneself or one's life experience—a process that is termed "overidentification." As Neff (2016a) writes, the various components of self-compassion are conceptually distinct and tap into different ways that individuals emotionally respond to pain and failure (with more kindness and less judgment), cognitively understand their predicament (as part of the human experience rather than as isolating), and pay attention to suffering (in a more mindful and less overidentified manner). The six elements of self-compassion are separable and do not covary in a lockstep manner, but they do mutually impact one another. Put another way, Neff (2016a, 2016b) proposes that self-compassion represents a dynamic system in which the various elements of self-compassion are in a state of synergistic interaction.

Over the last few years, research on self-compassion has grown at an exponential rate. There have been over 1,500 articles or dissertations written about self-compassion since 2003 (based on a Google Scholar search of entries with "self-compassion" in the title, May, 2018), over half of which have been published in the last 2 years. The majority of research studies have utilized the

supplementary materials with input from other authors. All other authors are listed in alphabetical order and contributed data as well as making comments on earlier drafts of the article. The clinical sample from the United Kingdom used in this study was drawn from the PREVENT trial, a project funded by the National Institute for Health Research Health Technology Assessment (NIHR HTA) Programme (project 08/56/01). This trial is reported in full in The

Lancet, [http://dx.doi.org/10.1016/S0140-6736\(14\)62222-4](http://dx.doi.org/10.1016/S0140-6736(14)62222-4). We are grateful to the trial team for allowing us to use the data.

Correspondence concerning this article should be addressed to Kristin D. Neff, Department of Educational Psychology, University of Texas at Austin, 1912 Speedway, Stop D5800, Austin, TX 78712-1289. E-mail: kristin.neff@mail.utexas.edu

Self-Compassion Scale (SCS; Neff, 2003a) to examine the construct of self-compassion. The SCS is intended to be used as a total score to measure self-compassion, or else as six subscale scores to assess its constituent elements: Neff (2016a, 2016b) proposes that the state of self-compassion entails more compassionate and fewer uncompassionate responses to personal suffering, which is why the SCS measures both.

Neff's operationalization of the SCS was based on compassion for others as broadly conceptualized in Buddhist philosophy (Neff, 2003b), although scores on the SCS have a relatively weak correlation with compassion for others (Neff & Pommier, 2013). This appears to be because most people have significantly more compassion for others than for themselves (Neff, 2003a; Neff & Pommier, 2013), meaning the two do not necessarily go hand in hand.

Research using the SCS suggests that self-compassion is a key indicator of wellbeing. For instance, cross-sectional research using the SCS shows that self-compassion has moderate to strong positive associations with outcomes such as happiness, optimism, life satisfaction, body appreciation and motivation, and negative associations with outcomes such as depression, anxiety, maladaptive perfectionism, and fear of failure—findings that are replicated using experimental methods such as interventions or mood manipulations (see Neff & Germer, 2017, for a review). While research suggests that self-compassion yields similar mental health benefits as other positive self-attitude constructs such as self-esteem (Neff, 2011), it does not appear to have the same pitfalls (Crocker & Park, 2004). For instance, Neff and Vonk (2009) found that while self-compassion and self-esteem were strongly correlated, simultaneous regressions indicated that self-compassion was associated with more stable and less contingent feelings of self-worth over time, and was associated with less social comparison, public self-consciousness, self-rumination, anger, closed-mindedness, and narcissism than self-esteem. Similarly, an experience sampling study conducted by Krieger, Hermann, Zimmermann, and Grosse Holtforth (2015) found that levels of self-compassion, but not self-esteem, predicted less negative affect when encountering stressful situations over a 14-day period.

The incremental predictive validity of SCS scores have been demonstrated with constructs such as neuroticism (Neff, Tóth-Király, & Colosimo, 2018; Stutts, Leary, Zeveney, & Hufnagle, in press) and self-criticism (Neff, 2003a). Although a key feature of self-compassion is the lack of self-judgment, overall SCS scores still negatively predict anxiety and depression when controlling for self-criticism and negative affect (Neff, Kirkpatrick, & Rude, 2007).

It should be mentioned that there are other models and measures of self-compassion, and that there is a lack of consensus in the field on how to define or measure compassion for self or others (Gilbert et al., 2017; Gilbert, Clarke, Hempel, Miles, & Irons, 2004; Strauss et al., 2016). Given that the SCS is the most commonly used measure of self-compassion, however, the current study is aimed at examining the psychometric properties of the SCS in a way that is theoretically consistent with Neff's (2003b) operationalization of the construct.

The SCS was developed in a sample of U.S. college undergraduates (Neff, 2003a). Confirmatory factor analyses (CFA) were used to provide support that scale items fit as intended with the a priori theoretical model (Furr & Bacharach, 2008). An initial CFA

found an adequate fit for a six-factor intercorrelated model and a higher-order factor model. Since that time, at least 30 published studies have examined the factor structure of the SCS (see Table 1 in the supplemental materials for a summary). Multiple translations of the SCS have been published, most of which have replicated the six-factor structure of the SCS using CFA. While not all examined the higher-order model, those that did yielded inconsistent findings. For example, a higher-order factor was supported with a Czech (Benda & Reichová, 2016), Norwegian (Dundas, Svendsen, Wiker, Granli, & Schanche, 2016), and two Portuguese samples (Castilho, Pinto-Gouveia, & Duarte, 2015; Cunha, Xavier, & Castilho, 2016), but not with German (Hupfeld & Ruffieux, 2011), Italian (Petrocchi, Ottaviani, & Couyoumdjian, 2013), or a third Portuguese sample (Costa et al., 2016).

Recently, there has been controversy over whether or not self-compassion should be measured as an overall construct, or if compassionate versus uncompassionate self-responding should be measured separately. Some have found that use of a total score is not justified through higher-order factor analyses, and have argued that two separate positive and negative factors demonstrate better fit (e.g., Costa et al., 2016; López et al., 2015; Montero-Marín et al., 2016). These researchers tend to use the term "self-compassion" to describe the positive factor and "self-criticism" or "self-coldness" to describe the negative factor (Costa et al., 2016; Gilbert, McEwan, Matos, & Rivis, 2011; López et al., 2015). However, self-criticism and self-coldness primarily describe self-judgment, or how people emotionally respond to suffering, and do not describe isolation (a way of cognitively understanding suffering) or overidentification (a way of paying attention to suffering). Moreover, these terms may obscure the fact that items representing negative self-responding are reverse-coded to indicate their relative absence. Therefore, we prefer the terms compassionate self-responding (CS) to represent the three components of self-kindness, common humanity and mindfulness and reduced uncompassionate self-responding (RUS) to represent lessened self-judgment, isolation, and overidentification measured by the SCS.

The question of whether the SCS can be used to measure self-compassion as a holistic state of being or if it should be used to measure two distinct states of being has important implications for our understanding of what self-compassion is. If self-compassion does not include RUS, the implication would be that how self-critical, isolated, and overidentified individuals are in times of struggle has little bearing on how self-compassionate they are. This, in turn, would have implications for researchers' attempts to examine the link between self-compassion and wellbeing. For instance, Muris and Petrocchi (2017) conducted a meta-analysis of the link of the SCS subscales with psychopathology across 18 studies and found the three components representing RUS had a stronger association with psychopathology (e.g., depression, anxiety, and stress) than the CS components. They argue that negative items "may inflate the relationship with psychopathology" (p. 734) and should, therefore, be excluded from the SCS. If, however, RUS is an integral part of self-compassion, then logically speaking it cannot "inflate" its own association with psychopathology. Rather, RUS could be interpreted to "explain" the link between self-compassion and psychopathology. Support for this point of view can be found in studies designed to examine self-compassion through mood induction (i.e., using writing prompts) or through intervention, which show that increasing

self-compassion experimentally also leads to reduced negative outcomes such as depression, anxiety, shame, and so forth (see Neff & Germer, 2017). Not including RUS subscales in the measurement of self-compassion, therefore, could potentially underestimate its relationship to psychopathology.

Some have argued that the CS and RUS subscales should not be combined into a total self-compassion score because compassionate responding is associated with parasympathetic nervous system activity and uncompassionate responding with sympathetic activity (Gilbert et al., 2011). However, research with the SCS shows that the CS and RUS subscales do not substantially differ in terms of their association with markers of sympathetic response (e.g., α -amylase, interleukin-6) after a stressful situation (Neff, Long et al., in press), or vagally mediated heart-rate variability, a marker of parasympathetic response (Svendsen et al., 2016). As Porges (2001) makes clear, the two types of autonomic nervous system responding themselves interact and covary as a system. The issue of whether self-compassion is best measured as a total score or if CS and RUS should be measured separately is largely a psychometric question, however, which has yet to be definitively established.

Alternative Models for Examining the Factor Structure of the SCS

It is important that the psychometric analyses used to examine psychological measures be consistent with the psychological theory underlying those measures (Morin, Arens, Tran, & Caci, 2016b). Higher-order models are commonly used to validate the simultaneous use of a total score and subscale scores in measures of multidimensional psychological constructs (e.g., Chen, West, & Sousa, 2006; Gignac, 2016). A higher-order model represents several first-order factors (representing subscale scores) and a higher-order factor (representing a total score) that explains their intercorrelation, but makes the strong assumption that the higher-order factor only influences individual item responses through the pathway of the first-order factors (appropriate for certain constructs like IQ). The original SCS publication (Neff, 2003a) used a higher-order model to justify use of a total and six subscale scores, but as mentioned above, support for a higher-order model has been inconsistent.

Williams, Dalgleish, Karl, and Kuyken (2014) did not find support for a higher-order factor in four different English samples (student, community, meditator, and clinical), but did find support for a six-factor correlated model. They suggested that the six subscales but not a total score be used. López et al. (2015) examined a Dutch community sample and did not find support for a higher-order factor, so conducted exploratory factor analysis (EFA) and found that the positive items loaded on one factor and the negative items loaded on a second factor. No CFA was conducted to confirm this two-factor model, however. Costa et al. (2016) examined a Portuguese clinical sample and compared a higher-order model, a six-factor uncorrelated model, a two-factor uncorrelated model that separated positive and negative items, and a two-factor model that included correlated errors designed to improve model fit, and found that the two-factor model with correlated errors had the best fit. These latter two sets of researchers suggested that separate positive and negative scores be used rather than a total score.

The bifactor model is an increasingly popular way to model multidimensional constructs (Reise, 2012; Rodriguez, Reise, & Haviland, 2016a). A bifactor model does not assume that the general or group factors are higher or lower than the other but rather coexist, and models the direct association of the general factor and group factors on individual item responses. This has the added benefit of enabling the calculation of omega values (ω) that represent the amount of reliable variance in item responding explained by the general factor. Note that with a bifactor model the group factors are not allowed to correlate. Although perhaps counterintuitive, this improves interpretability. For instance, it models those aspects of an item (e.g., When something upsets me I try to keep my emotions in balance) that are shared by all items in the general factor (e.g., self-compassion), as well as those aspects that are only shared by other items in its group factor (e.g., mindfulness). Neff (2016a) argued that a bifactor model provides a better theoretical fit with her conceptualization of self-compassion than a higher-order model given that behaviors assessed by individual items are directly representative of self-compassion as a general construct in addition to its constituent group components.

Neff, Whittaker, and Karl (2017) examined the SCS using bifactor CFA analysis in four different U.S. populations: undergraduates, community adults, meditators, and a clinical population. While the one-factor, two-factor correlated, and higher-order models had poor fit across samples, the six-factor correlated and bifactor models had acceptable fit using liberal fit criteria in the undergraduate, community, and meditator samples. Fit was inadequate in the clinical sample. Nonetheless, ω values revealed that over 90% of the reliable variance in scores could be explained by a general self-compassion factor in all four populations (including the clinical sample). Findings were interpreted as providing support for use of a total score as well as six subscale scores, but not as two positive and negative scores. Cleare, Gumley, Cleare, and O'Connor (2018) independently replicated these findings in a Scottish sample: support was not found for a one-factor, two-factor, or higher-order model, but was found for a six-factor correlated and bifactor model, with 94% of the variance in item responding explained by a general factor.

Three additional studies on translations of the SCS have provided evidence for a model with six group factors and one general factor using a bifactor CFA approach: French (Kotsou & Leys, 2016), Brazilian Portuguese (Souza & Hutz, 2016), and Italian (Veneziani, Fuochi, & Voci, 2017). However, Montero-Marín et al. (2016) did not find support for a bifactor CFA model in two Spanish and Brazilian-Portuguese samples of doctors, but did find support for two higher-order factors (CS and RUS) and six first-order factors. Moreover, Brenner, Heath, Vogel, and Credé (2017) found that a two-bifactor CFA model with six group factors and two uncorrelated general (CS and RUS) factors had better fit than a single-bifactor model in a sample of U.S. undergraduates, though findings for some indicators were poor and the choice of examining two uncorrelated general factors is not theoretically consistent with the construct of self-compassion. Thus, the dimensionality of the SCS is still in question. Also, the above-mentioned results suggest that the assumptions of CFA might be overly restrictive for the SCS, given the inconsistency of findings.

CFA makes the strict assumption that items can only load on their respective factors, and may fail to account for two main sources of construct-relevant dimensionality in complex scales

like the SCS, potentially resulting in biased parameters (Morin, Arens, & Marsh, 2016a; Morin, Arens, et al., 2016b). These sources do not refer to random measurement error, but are related to the idea that items often present more than one source of true score variance and subsequently belong to more than one construct. The first source refers to the fact that individual items are expected to be associated with a global factor (e.g., self-compassion), in which specific factors are not differentiated, as well as specific group factors (e.g., self-kindness or reduced self-judgment), in which they are differentiated. As mentioned, the relation between specific and global factors can be modeled in a hierarchical or in a bifactor manner with the latter generally being preferred unless there are strong theoretical reasons for the application of the former.

The second source of dimensionality comes from the fact that the six components of the scale are conceptually close and interrelated as a system, meaning items within each subscale should be expected to have significant associations with items in other subscales. Indeed, a recent review of simulation studies (Asparouhov, Muthén, & Morin, 2015) have shown that when cross-loadings between items and nontarget factors are not expressed (i.e., cross-loadings are constrained to be zero), parameters are likely to be biased. Exploratory Structural Equation Modeling (ESEM) is specifically designed to model system level interactions (Marsh, Morin, Parker, & Kaur, 2014; Morin, Marsh, & Nagengast, 2013). In CFA, items are strictly allowed to load on one factor, and additional associations between items and nontarget factors are reflected in the form of modification indices and/or inflated interfactor correlations, which are the only ways overlap can be expressed. In ESEM, these associations are expressed in the form of item cross-loadings. Unlike EFA, in which no a priori hypotheses about models are advanced, ESEM with target rotation (Browne, 2001) can model a priori hypotheses and, therefore, be directly compared with CFA models (Marsh et al., 2014). ESEM has been suggested to result in substantially better fit and less strongly correlated factors than corresponding CFA solutions (Marsh, Liem, Martin, Morin, & Nagengast, 2011; Morin & Maïano, 2011; Tóth-Király, Orosz, et al., 2017).

ESEM has rarely been used to examine the SCS. However, Hupfeld and Ruffieux (2011) as well as Tóth-Király, Bóthe, and Orosz (2017) applied ESEM to analyze the factor structure of the SCS and found that, compared with CFA, ESEM provided a better fit to the data. Moreover, to account for the two sources of construct-relevant dimensionality, Tóth-Király et al. (2017) used the integrative bifactor ESEM framework (Morin, Arens, et al., 2016a, 2016b; Morin, Boudrias, Marsh, Madore, & Desrumaux, 2016), and results strongly supported the presence of a global self-compassion factor as well as six specific factors. The overarching bifactor ESEM framework appears to be especially appropriate for the SCS because it can simultaneously model both the specific and overall relationship of items using a bifactor analytic approach as well as their interaction as a system with an ESEM approach.

The Current Study

In the current study, we examined the factor structure of the SCS using both CFA and ESEM analyses for five distinct models: a

single factor, two-factor correlated, six-factor correlated, single-bifactor model (one general factor and six group factors), and a correlated two-bifactor model (a general factor representing CS with three group factors representing higher levels of self-kindness, common humanity and mindfulness, and a general factor representing RUS with three group factors representing lower levels of self-judgment, isolation, and overidentification). Based on the existing literature, we expected that the one-factor and two-factor correlated models would have poor fit, and the six factor-correlated, single-bifactor, and two-bifactor models would have better fit. We also expected fit indices to be better in ESEM rather than CFA analyses given that it is more appropriate for modeling system-level interactions. Our overall goal was to determine the best-fitting solution that is also well-aligned with Neff's underlying model of self-compassion (Neff, 2003b), given that this is the theoretical model that the SCS was created to measure.

We examined the SCS in 20 different samples. Because the SCS was developed in English we included 7 English samples, but also 13 samples from non-English speaking countries. We included student, community, meditator, and clinical samples. The meditator and one of the clinical samples were the same as examined in Neff et al. (2017), and a second Portuguese clinical sample was also included (Castilho et al., 2015). Given that the SCS is commonly used to assess outcomes of meditation-based and clinical interventions (e.g., Birnie, Speca, & Carlson, 2010; Kelly, Wisniewski, Martin-Wagar, & Hoffman, 2017), we felt it was important to include these populations. The comprehensiveness of this study was designed to try to find more definitive answers to questions regarding the factor structure of the SCS: Should a total score be used, or two separate scores representing CS and RUS?

Method

Procedure

This study was organized by the first three authors, who wanted to examine the factor structure of the SCS in a variety of international samples. SCS data for three samples from the United States were contributed by the first and third authors, who originally collected the data for other research purposes. Appropriate Institutional Review Board (IRB) approval was received before collecting these data, which were de-identified for the current study before being statistically analyzed by the second author. To gather SCS data from samples outside of the United States, researchers were contacted in other English and non-English-speaking countries. These researchers contributed SCS data for 17 additional samples, which had also been collected previously for other research purposes (Information about the data source of each sample as well as participant recruitment procedures can be found in the online supplementary materials.). SCS data contributed from sources outside the United States were received as secondary data and included no potential participant identifiers. IRB approval was not required for analyses of these de-identified secondary data, although researchers from outside the United States also obtained local ethics committee approval before collecting their original data as appropriate.

Participants

The initial number of participants was 11,990 from 20 international samples drawn from the following countries: Australia, Brazil, Canada, China, France, Germany, Greece, Iran, Italy, Japan, South Korea, Norway, Portugal, Spain, United Kingdom, and United States. In total, we included 10 community, 6 student, 1 mixed community/student, 1 meditator, and 2 clinical samples. Participants were excluded if they were under age 18 or had more than 50% of their responses missing. Thus, the final sample included 11,685 respondents (3,296 males, 8,367 females, 22 unspecified), aged between 18 and 83 ($M = 32.29$, $SD = 8.28$). Specific sample characteristics can be seen in Table 1.

Measures

The SCS (Neff, 2003a) is a 26-item self-report questionnaire measuring the six components of self-compassion: Self-Kindness (5 items; e.g., “I try to be loving towards myself when I’m feeling emotional pain”), reduced Self-Judgment (5 items; e.g., “I’m disapproving and judgmental about my own flaws and inadequacies”), Common Humanity (4 items; e.g., “When things are going badly for me, I see the difficulties as part of life that everyone goes through”), reduced Isolation (4 items; 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”), Mindfulness (4 items; e.g., “When I’m feeling down I try to approach my feelings with curiosity and openness”), and reduced Overidentification (4 items; e.g., “When something upsets me I get carried away with my feelings”). Responses are given on a scale from 1 (*almost never*) to 5 (*almost always*). Note that all items in the Self-Kindness, Common Humanity, and Mindfulness subscales are positively valenced, while all items in the Self-Judgment, Isolation, and Overidentification subscales are negatively valenced. Items representing uncom-

passionate self-responding are reverse-coded before calculating a total score to indicate their relative absence in a self-compassionate mindset. Means are calculated for each subscale, and a grand mean is calculated for a total self-compassion score. Neff (2003a) found that items forming a total SCS score evidenced good internal reliability (Cronbach’s $\alpha = .92$), as did the six subscales (Cronbach’s α ranging from .75 to .81). Test-retest reliability over a 3-week interval was good (total score, Cronbach’s $\alpha = .93$; six subscales, Cronbach’s α ranging from .80 to .88). The current study also used 12 SCS translations (out of 16 published): Brazilian Portuguese, Chinese, French, German, Greek, Persian, Italian, Japanese, Korean, Norwegian, Portuguese, and Spanish. A description of the psychometric properties of each SCS translation can be found in the online supplementary materials.

Analyses

All statistical analyses were conducted with Mplus 7.4 (Muthén & Muthén, 1998–2017) with the weighted least squares mean- and variance-adjusted estimator (WLSMV) as it is more suitable for ordered-categorical items with five or less answer options than estimators based on maximum-likelihood (e.g., Bandalos, 2014; Finney & DiStefano, 2006). Before the main analyses, negative items were reverse-coded. To systematically investigate the potential sources of construct-relevant dimensionality of the SCS, five corresponding CFA and ESEM models were tested and subsequently contrasted: (1a, 1b) a one-factor model with a unitary self-compassion dimension; (2a, 2b) a two-factor correlated model with two unitary factors representing CS and RUS; (3a, 3b) a six-factor correlated model with six components of self-compassion; (4a, 4b) a single-bifactor model with a general self-compassion factor and six group factors; and (5a, 5b) a two-bifactor model including two correlated general factors representing CS and RUS, each with three group factors. As per

Table 1
Characteristics of the Total and Individual Samples

Country	Language	Type	Initial <i>N</i>	Final <i>N</i>	Females	<i>M</i> age (<i>SD</i>)
Total		Combined	11,990	11,685	8,367 (71.6%)	32.29 (8.28)
AUS	English	Community	316	316	240 (75.9%)	37.20 (14.67)
BRA	Brazilian Portuguese	Community	312	312	241 (77.2%)	30.36 (10.76)
CAN	English	Student	395	362	308 (85.1%)	21.23 (4.02)
CHI	Chinese	Community	262	261	255 (97.7%)	37.02 (7.68)
FRA	French	Community	1,554	1,545	1,362 (88.2%)	43.07 (12.48)
GER	German	Community	396	380	303 (79.7%)	29.43 (10.15)
GRE	Greek	Community	981	974	612 (62.8%)	21.99 (6.09)
IRA	Persian	Student	575	448	239 (53.3%)	25.33 (7.38)
ITA	Italian	Community	384	380	257 (67.6%)	33.56 (10.46)
JAP	Japanese	Student	718	718	291 (40.5%)	19.42 (1.16)
KOR	Korean	Student	353	343	180 (52.5%)	38.80 (9.22)
NOR	Norwegian	Student	327	318	189 (59.4%)	23.03 (3.40)
POR 1	Portuguese	Mixed	1,128	1,101	824 (74.8%)	24.71 (8.01)
POR 2	Portuguese	Clinical	314	297	236 (79.5%)	29.37 (8.43)
SPA	Spanish	Community	434	434	306 (70.5%)	49.71 (10.83)
UK 1	English	Community	1,108	1,085	969 (89.3%)	21.38 (5.69)
UK 2	English	Clinical	390	390	300 (76.9%)	50.16 (11.08)
US 1	English	Community	984	974	619 (63.6%)	38.17 (12.88)
US 2	English	Student	844	833	486 (58.3%)	21.22 (3.53)
US 3	English	Meditator	215	214	150 (70.1%)	47.36 (11.62)

Note. AUS = Australia; BRA = Brazil; CAN = Canada; CHI = China; FRA = France; GER = Germany; GRE = Greece; IRA = Iran; ITA = Italy; JAP = Japan; KOR = South Korea; NOR = Norway; POR = Portugal; SPA = Spain; UK = United Kingdom; US = United States.

typical model specifications, in the CFA-based models (1a–5a), items were only allowed to load on their a priori target factors with cross-loadings being constrained to zero. In the ESEM-based models (1b–5b), items were allowed to load on the nontarget factors as well. ESEM was also estimated in a confirmatory manner with target rotation (Browne, 2001) as per prior suggestions (Asparouhov & Muthén, 2009) and applications (Tóth-Király, Bóthe, Rigó, & Orosz, 2017). In the correlated models (2a, 2b, 3a, and 3b), factors were allowed to correlate freely. In the case of the bifactor models (4a, 4b, 5a, and 5b), group factors were specified as orthogonal to the general factor, as is standard (e.g., Reise, 2012; Reise, Moore, & Haviland, 2010) but the two general factors were specified as correlated¹ (see also Tóth-Király, Morin, Bóthe, Orosz, & Rigó, 2018 for a similar application or Morin, Myers, & Lee, in press, for an overview). These models were tested in the total sample and individual samples.

In model assessment, instead of only relying on the χ^2 test that is sensitive to sample-size (Marsh, Hau, & Grayson, 2005), commonly applied goodness-of-fit indices were examined with their respective thresholds (Hu & Bentler, 1999; Marsh et al., 2005): the comparative fit index (CFI; $\geq .95$ for good, $\geq .90$ for acceptable), the Tucker–Lewis index (TLI; $\geq .95$ for good, $\geq .90$ for acceptable), the root-mean-square error of approximation (RMSEA; $\leq .06$ for good, $\leq .08$ for acceptable) with its 90% confidence interval, and the weighted root-mean-square residual (WRMR; ≤ 1.00 for acceptable). Note that we did not compare fit using Akaike’s Information Criterion (AIC) or Bayesian Information Criterion (BIC) because these information criteria are only available for maximum likelihood-based estimations, which are less accurate for ordered categorical data. However, the primary purpose of these indices is to determine which models would be most likely to cross-validate in subsequent samples, and this study determines cross-validation directly by examining model fit in 20 different samples.

Analyses of data should not be based solely on fit indices. The close inspection of parameter estimates (e.g., factor loadings, cross-loadings, and interfactor correlations) and the theoretical conformity of the models may also reveal valuable information about measurement models (as proposed by Hu & Bentler, 1998; Marsh, Hau, & Wen, 2004; Marsh et al., 2011; Morin, Arens, et al., 2016a, 2016b; Morin, Boudrias, et al., 2016). When comparing first-order CFA and ESEM models, the emphasis should be on comparison of factor correlations and on the need to incorporate cross-loadings, assuming that both solutions have well-defined factors with strong target loadings. If there is a substantial difference in the size of correlations between CFA and ESEM, the latter results are preferred as they provide more exact estimates (Asparouhov et al., 2015). If differences are negligible, then CFA is preferred because of its greater parsimony. Relatively large cross-loadings in the ESEM model may suggest an unmodeled general factor, which can be tested with a bifactor model. The general factor should also be well-defined by strong and theoretically meaningful factor loadings. Additionally, reduced cross-loadings and some well-defined specific factors would also provide support for the bifactor representation.² A particularly important question relates to the inclusion of one or two general factors where, once again, the close examination of factor loadings is highly informative.

We also assessed the reliability of items in the models. In the case of the six-factor model we calculated composite reliability (Raykov, 1997) as opposed to Cronbach’s α , which has been criticized as being less useful for determining the reliability of factors (Rodriguez et al., 2016a). It has the advantage of being model-based, taking into account factor loadings and item-specific measurement errors as well. Based on Bagozzi and Yi (1988), values above .60 are considered acceptable, whereas values above .70 are good. As bifactor models allow the partitioning of the different sources of variance into the global and specific factors, omega (ω) and omega hierarchical (ω_H) were also calculated for the best fitting models based on standardized estimates (Brunner, Nagy, & Wilhelm, 2012; Rodriguez, Reise, & Haviland, 2016b). Omega estimates the proportion of the variance in item responding that is attributed to both the global and specific factors. OmegaH estimates the proportion of variance that is attributed to the general factor only. Finally, we also compared the ω and ω_H s on the basis of Rodriguez et al. (2016b) to investigate the degree of reliable variance in item responding. For the variance attributed to the general factor, one should divide the value of ω_H by ω (i.e., ω_H/ω); for the remaining reliable variance attributable to the specific factors, one should subtract ω_H from ω (i.e., $\omega - \omega_H$). Reise, Bonifay, and Haviland (2013) suggest 75% or higher accounted for by the general factor as the ideal amount of variance to justify use of a total score despite the presence of multidimensionality in the data.

Results

Structural Analyses

Because results were generally similar for the total sample and the individual samples, we mainly refer to results for the total sample for the sake of simplicity. We first examined the fit of the one-factor model for all samples (see the online supplementary materials). The one-factor ESEM solution is fundamentally a one-factor CFA (using only different estimation routines in Mplus) as there are no cross-loadings in this model. In accord with our hypotheses, results clearly demonstrate the inadequacy of the unidimensional model (Total sample: CFI = .74, TLI = .73, RMSEA = .15 [90% confidence interval, CI .15–.15], WRMR = 14.44). Tables 2, 3, 4, and 5 present model fit indices for CFA and ESEM analyses for the two-factor, six-factor, single-bifactor, and correlated two-bifactor models, respectively. In the case of the two-factor correlated models (see Table 2), both the CFA (Total sample: CFI = .90, TLI = .89, RMSEA = .10 90% CI [.09–.10], WRMR = 7.48) and ESEM (Total sample: CFI = .88, TLI = .86, RMSEA = .11 90% CI [.11–.11], WRMR = 6.31) versions showed marginally acceptable fit indices in some samples, but the

¹ In the two-bifactor Exploratory Structural Equation Modeling (ESEM) model, the general factors were specified as confirmatory factor analysis (CFA) factors (i.e., no cross-loadings between them), while the six specific factors were specified as ESEM factors (i.e., cross-loadings between them were allowed).

² Naturally, not all specific factors are well-defined in the bifactor model relative to the first-order model, given that the item-level covariance is disaggregated to two sources (general and specific factors) instead of one (e.g., Morin et al., 2016a; Tóth-Király et al., 2018).

Table 2
Goodness-of-Fit Indices for the Total and Individual Samples: Two-Factor Correlated Models

Country	Two-factor correlated CFA					Two-factor correlated ESEM				
	CFI	TLI	RMSEA	90% CI	WRMR	CFI	TLI	RMSEA	90% CI	WRMR
Total	.90	.89	.10	[.09–.10]	7.48	.88	.86	.11	[.11–.11]	6.31
AUS	.93	.92	.10	[.09–.10]	1.51	.92	.91	.10	[.10–.11]	1.25
BRA	.94	.93	.08	[.08–.09]	1.36	.94	.93	.08	[.07–.09]	1.13
CAN	.89	.88	.09	[.09–.10]	1.70	.89	.87	.10	[.09–.10]	1.38
CHI	.96	.96	.10	[.10–.11]	1.41	.96	.95	.11	[.10–.12]	1.20
FRA	.89	.88	.11	[.11–.11]	3.18	.89	.87	.12	[.11–.12]	2.68
GER	.81	.80	.11	[.10–.12]	1.89	.84	.81	.10	[.10–.11]	1.51
GRE	.92	.92	.08	[.08–.09]	2.19	.90	.88	.10	[.10–.10]	1.91
IRA	.81	.79	.09	[.08–.09]	1.88	.90	.88	.07	[.06–.07]	1.19
ITA	.85	.84	.12	[.12–.13]	2.15	.87	.84	.12	[.11–.12]	1.62
JAP	.86	.84	.10	[.09–.10]	2.64	.78	.73	.13	[.12–.13]	2.52
KOR	.91	.90	.09	[.08–.10]	1.79	.94	.93	.08	[.07–.09]	1.08
NOR	.89	.88	.10	[.09–.10]	1.68	.88	.86	.10	[.10–.11]	1.47
POR 1	.92	.92	.10	[.10–.10]	2.75	.90	.89	.12	[.11–.12]	2.22
POR 2	.89	.88	.10	[.09–.10]	1.57	.89	.87	.10	[.09–.11]	1.37
SPA	.82	.80	.11	[.11–.11]	2.25	.88	.85	.09	[.09–.10]	1.43
UK 1	.88	.87	.11	[.10–.11]	2.68	.88	.85	.11	[.11–.11]	2.27
UK 2	.89	.88	.09	[.08–.09]	1.68	.88	.85	.10	[.09–.10]	1.40
US 1	.91	.90	.10	[.09–.10]	2.29	.92	.90	.10	[.09–.10]	1.86
US 2	.83	.81	.11	[.10–.11]	2.61	.86	.83	.10	[.10–.10]	1.93
US 3	.92	.91	.10	[.09–.11]	1.37	.92	.91	.10	[.10–.11]	1.16

Note. CFA = confirmatory factor analysis; ESEM = exploratory structural equation modeling; CFI = comparative fit index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; WRMR = weighted root mean square residual; CI = confidence interval; AUS = Australia; BRA = Brazil; CAN = Canada; CHI = China; FRA = France; GER = Germany; GRE = Greece; IRA = Iran; ITA = Italy; JAP = Japan; KOR = South Korea; NOR = Norway; POR = Portugal; SPA = Spain; UK = United Kingdom; US = United States.

majority were not acceptable by commonly applied standards, hence we rejected these solutions. In the case of the six-factor correlated CFA and ESEM models (see Table 3), most CFA models had acceptable fit (Total sample: CFI = .95, TLI = .94, RMSEA = .07 [90% CI .07–.07], WRMR = 5.15). However, ESEM systematically outperformed these solutions as apparent with excellent fit indices (Total sample: CFI = .99, TLI = .97, RMSEA = .05 [90% CI .05–.05], WRMR = 1.75).

Following Morin et al. (2016a, 2016b), we also examined standardized item factor loadings for the corresponding CFA and ESEM solutions for the total sample to select the final models, presented in Tables 6, 7, and 8. When examining the six-factor correlated models (see Table 6), all six factors were well-defined by their respective factor loadings ($\lambda = .65$ to $.84$, $M_\lambda = .76$) in CFA, but this solution also resulted in relatively high factor correlations ($r = .38$ to $.91$, $Mr = .64$), undermining the discriminant validity of interpretations of items in the six factors. In the ESEM model, factor loadings ($\lambda = .26$ to $.97$, $M_\lambda = .56$) as well as factor correlations ($r = .16$ to $.66$, $Mr = .42$) were systematically lower. These results are in line with previous studies (Morin et al., 2016a) showing that ESEM often provides a better representation of the interfactor correlations. As expected, there were some cross-loadings ($|\lambda| = .00$ to $.42$, $M_\lambda = .10$) between conceptually similar items ($\geq .32$; Worthington & Whittaker, 2006). For example, the self-kindness item “I’m tolerant of my own flaws and inadequacies” cross-loaded on reduced self-judgment. Overall, cross-loadings were found for 2 self-kindness items on reduced self-judgment, 1 self-kindness item on mindfulness, 1 reduced self-judgment item on self-kindness, 1 mindfulness item on self-kindness, 1 mindfulness item on on reduced overidentification, and 2 reduced overidentification items on reduced self-judgment.

The next question that we addressed is whether the single-bifactor model with one general factor (representing self-compassion) or the correlated two-bifactor model with two general factors (representing CS and RUS) was able to provide an improved representation of the data. For the single-bifactor models (see Table 4), CFA models were generally inadequate (Total sample: CFI = .85, TLI = .82, RMSEA = .12 90% CI [.12–.12], WRMR = 10.55), whereas ESEM models generally had much better fit (Total sample: CFI = .99, TLI = .98, RMSEA = .04 90% CI [.04–.04], WRMR = 1.42). Results for CFA and ESEM were less differentiated for the correlated two-bifactor models including two general factors (see Table 5), with generally adequate fit for the CFA models (Total sample: CFI = .96, TLI = .95, RMSEA = .06 90% CI [.06–.06], WRMR = 4.49), as well as for the ESEM models (Total sample: CFI = .99, TLI = .99, RMSEA = .04 90% CI [.03–.04], WRMR = 1.20). However, it should be noted that the correlated two-bifactor CFA model for the total sample had misspecifications, and almost half (9 out of 20) of the individual samples had negative residual variances, suggesting that the data did not support the hypothesized models. Therefore, we only compared the parameter estimates of the competing single- and correlated two-bifactor ESEM models.

The parameter estimates for the single-bifactor model (see Table 7) revealed a well-defined general factor ($|\lambda| = .36$ to $.75$, $M = .62$) reflecting a global level of self-compassion. As for the specific factors, common humanity retained a higher degree of specificity ($|\lambda| = .35$ to $.73$, $M = .53$) once the general factor was extracted. By the same token, isolation ($|\lambda| = .24$ to $.58$, $M = .41$) and mindfulness ($|\lambda| = .28$ to $.52$, $M = .41$) had moderate degree of specificity, self-kindness ($|\lambda| = .06$ to $.56$, $M = .34$) and overidentification ($|\lambda| = .19$ to $.50$, $M = .34$) had a smaller degree

Table 3
Goodness-of-Fit Indices for the Total and Individual Samples: Six-Factor Correlated Models

Country	Six-factor correlated CFA					Six-factor correlated ESEM				
	CFI	TLI	RMSEA	90% CI	WRMR	CFI	TLI	RMSEA	90% CI	WRMR
Total	.95	.94	.07	[.07-.07]	5.15	.99	.97	.05	[.05-.05]	1.75
AUS	.94 ^a	.93	.09	[.08-.10]	1.27	.98	.97	.06	[.06-.07]	.52
BRA	.96	.96	.06	[.06-.07]	1.05	.99	.98	.05	[.04-.06]	.51
CAN	.93	.92	.08	[.07-.08]	1.30	.97	.94	.06	[.06-.07]	.60
CHI	.97	.97	.09	[.08-.10]	1.17	.99 ^a	.99	.06	[.05-.07]	.46
FRA	.92	.91	.09	[.09-.10]	2.54	.98	.96	.06	[.06-.06]	.89
GER	.87 ^a	.85	.09	[.09-.10]	1.53	.98	.97	.05	[.04-.05]	.50
GRE	.97	.96	.06	[.05-.06]	1.40	.98	.97	.05	[.05-.06]	.68
IRA	.85 ^a	.83	.08	[.07-.08]	1.62	.96	.93	.05	[.04-.06]	.66
ITA	.91	.90	.10	[.09-.10]	1.60	.98	.97	.06	[.05-.06]	.53
JAP	.93	.92	.07	[.07-.07]	1.75	.96	.93	.06	[.06-.07]	.84
KOR	.92	.91	.09	[.08-.09]	1.60	.98	.96	.06	[.05-.06]	.53
NOR	.93	.92	.08	[.07-.08]	1.26	.98	.97	.05	[.04-.06]	.48
POR 1	.94 ^a	.94	.09	[.08-.09]	2.20	.99	.97	.06	[.05-.06]	.71
POR 2	.92 ^a	.91	.08	[.08-.09]	1.30	.97	.95	.06	[.05-.07]	.56
SPA	.86 ^a	.84	.10	[.09-.10]	1.90	.97	.95	.05	[.05-.06]	.58
UK 1	.94	.93	.08	[.07-.08]	1.80	.98	.97	.05	[.04-.05]	.67
UK 2	.92	.90	.08	[.07-.08]	1.41	.98	.96	.05	[.04-.06]	.55
US 1	.96	.95	.07	[.07-.07]	1.51	.99	.98	.04	[.04-.05]	.57
US 2	.92	.91	.07	[.07-.08]	1.73	.98	.96	.05	[.05-.06]	.67
US 3	.95	.95	.08	[.07-.09]	1.05	.99	.98	.05	[.04-.06]	.43

Note. CFA = confirmatory factor analysis; ESEM = exploratory structural equation modeling; CFI = comparative fit index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; WRMR = weighted root mean square residual; CI = confidence interval; AUS = Australia; BRA = Brazil; CAN = Canada; CHI = China; FRA = France; GER = Germany; GRE = Greece; IRA = Iran; ITA = Italy; JAP = Japan; KOR = South Korea; NOR = Norway; POR = Portugal; SPA = Spain; UK = United Kingdom; US = United States.

^a These solutions had model identification issues, suggesting overparameterization.

of specificity, whereas self-judgment ($|\lambda| = .07$ to $.44$, $M = .22$) retained almost no meaningful specificity. Finally, cross-loadings also slightly decreased in magnitude ($|\lambda| = .01$ to $.34$, $M = .09$) relative to the six-factor ESEM model. In the case of the correlated two-bifactor-ESEM model (see Table 8), while the correlation between the two factors was reduced, $r = .09$, $p = .086$, the two general factors were weakly defined by their respective factor loadings (Positive: $|\lambda| = .01$ to $.48$, $M = .22$; Negative: $|\lambda| = .04$ to $.35$, $M = .17$), arguing against the incorporation of a second general factor and supporting the superiority of the single-bifactor ESEM model with one general factor. Taking these results together, it appears that a six-factor correlated model (representing the six components of self-compassion) and a single-bifactor model (representing a general self-compassion factor and six specific factors) are supported, but a correlated two-bifactor model (representing CS and RUS) is not supported once parameter estimates are taken into account.

Reliability Analyses

Finally, we estimated composite reliability indices for items in the six-factor model and the ω and ω_H indices for items in the single-bifactor ESEM model to examine reliability. For the six-factor model (examining the sample as a whole), items in all factors had acceptable levels of composite reliability using Bagozzi and Yi's (1988) criteria of $>.60$ as adequate and $>.70$ as good: (self-kindness = $.84$, reduced self-judgment = $.73$, common humanity = $.81$, reduced isolation = $.83$, mindfulness = $.67$, and reduced overidentification = $.70$). (Composite reliability for items in the individual samples are available upon request from the first

author.) Reliability results for the single-bifactor model for all samples are presented in Table 9, although again we only discuss results for the total sample here. The single-bifactor model displayed high ω ($.96$) and ω_H ($.91$) values, demonstrating that a large majority of the variance in item responding can be attributed to the general factor. As per Rodriguez et al. (2016b), we compared the ratio of ω and ω_H to establish the amount of reliable variance of items attributable to the general factor (ω_H divided by ω) and that attributable to the multidimensionality caused by the specific factors (ω_H subtracted from ω). For the single-bifactor model, 95% of the reliable variance in item responding was attributed to the general self-compassion factor, whereas 5% was attributed to the group factors.

Discussion

Our analyses, which were designed to determine the best factor structure for the SCS, found that a one- and two-factor solution to the SCS had an inadequate fit using both CFA and ESEM. In contrast, a six-factor correlated solution had good fit using ESEM (CFA results for the six-factor solution were also acceptable) in every sample examined. The single-bifactor ESEM model (with one group and six specific factors) also had good fit in every sample. Moreover, inspection of factor loadings suggested good parameter estimates for a single general factor in ESEM. While the correlated two-bifactor ESEM model with two correlated general factors also had good model fit, factor loadings indicated poor specification of separate factors representing CS and RUS, so this model was rejected. Note that the single-bifactor ESEM model also had the highest level of theoretical conformity with Neff's

Table 4
Goodness-of-Fit Indices for the Total and Individual Samples: Bifactor Models

Country	Bifactor CFA					Bifactor ESEM				
	CFI	TLI	RMSEA	90% CI	WRMR	CFI	TLI	RMSEA	90% CI	WRMR
Total	.85	.82	.12	[.12-.12]	10.55	.99	.98	.04	[.04-.04]	1.42
AUS	.92	.90	.11	[.10-.11]	1.53	.99	.98	.05	[.04-.06]	.40
BRA	.93	.92	.09	[.08-.09]	1.34	.99	.98	.05	[.04-.06]	.45
CAN	.85	.82	.12	[.11-.12]	2.31	.97	.95	.06	[.05-.07]	.51
CHI	.95 ^a	.94	.12	[.11-.13]	1.57	.99 ^a	.99	.05	[.04-.06]	.39
FRA	.89	.87	.11	[.11-.12]	3.23	.99	.98	.05	[.05-.05]	.69
GER	.88	.85	.09	[.09-.10]	1.53	.99	.97	.04	[.03-.05]	.43
GRE	.83	.80	.13	[.13-.13]	3.51	.99	.98	.04	[.04-.05]	.53
IRA	.67	.61	.12	[.12-.13]	2.34	.97	.94	.05	[.04-.06]	.57
ITA	.89	.87	.11	[.10-.11]	1.88	.99	.97	.05	[.04-.06]	.45
JAP			No identification			.97 ^a	.95	.06	[.05-.06]	.68
KOR	.63	.56	.19	[.19-.20]	3.96	.98	.97	.05	[.04-.06]	.45
NOR	.87	.85	.11	[.10-.11]	1.70	.99	.97	.05	[.04-.06]	.43
POR 1	.83	.80	.15	[.15-.15]	4.61	.99	.98	.05	[.05-.06]	.62
POR 2	.85	.82	.12	[.11-.12]	1.79	.98	.95	.06	[.05-.07]	.48
SPA	.74	.69	.14	[.13-.14]	2.84	.98	.96	.05	[.04-.06]	.48
UK 1	.89	.87	.11	[.10-.11]	2.73	.99	.98	.04	[.04-.05]	.55
UK 2	.82 ^a	.79	.11	[.11-.12]	2.04	.98	.97	.05	[.04-.05]	.48
US 1	.90	.88	.11	[.10-.11]	2.54	.99	.99	.04	[.03-.04]	.49
US 2	.80	.76	.12	[.11-.12]	3.20	.98	.96	.05	[.04-.05]	.57
US 3	.92	.90	.11	[.10-.11]	1.32	.99	.99	.04	[.02-.05]	.36

Note. CFA = confirmatory factor analysis; ESEM = exploratory structural equation modeling; CFI = comparative fit index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; WRMR = weighted root mean square residual; CI = confidence interval; AUS = Australia; BRA = Brazil; CAN = Canada; CHI = China; FRA = France; GER = Germany; GRE = Greece; IRA = Iran; ITA = Italy; JAP = Japan; KOR = South Korea; NOR = Norway; POR = Portugal; SPA = Spain; UK = United Kingdom; US = United States.

^a These solutions had model identification issues, suggesting overparameterization.

(2003b) view that self-compassion is comprised of six components that interact as a global system. Results for our final selected models were remarkably similar across the 20 diverse populations examined—including student, community, clinical, and meditator samples in 13 different languages—providing strong support for the generalizability of the SCS to measure self-compassion.

Findings regarding cross-loadings in the ESEM models are also informative. In the six-factor model all factors were well defined, but eight cross-loadings were found (cross loadings were found equally within and across the CS and RUS dimensions). These cross-loadings highlight the importance of using models such as ESEM that can uncover this particular source of construct-relevant dimensionality. Use of a total SCS score was supported by the finding that 95% of the reliable variance in SCS item responding could be explained by a general factor for the total sample, ranging from 86 to 96% for the individual samples. This is well over the 75% or higher suggested by Reise et al. (2013) to justify use of a total score. All of the factors in the six-factor solution had adequate to good levels of composite reliability based on conventional thresholds (Bagozzi & Yi, 1988). In the single-bifactor model the general factor was well-defined and the specific factors were moderately well-defined. These observations give support for the idea that the specific factors assess relevant components over and above the general factors. They also support the system level interaction of components. We interpret these results as supporting use of a global score (representing self-compassion) or six subscale scores (representing self-kindness, common humanity, mindfulness and reduced self-judgment, isolation and overidentification), but not two separate CS and RUS scores.

The fact that the one- and two-factor solution had poor fit but a six-factor solution had good fit makes sense theoretically. It is potentially problematic to argue that self-compassion is a unitary construct (no theorists we are aware of have made this argument), or to argue that the three subscales representing CS versus RUS each form unitary constructs, as proposed by some (e.g., Costa et al., 2016; López et al., 2015). The three subscales within each of these dimensions are distinct, and tap into the way that people *emotionally* respond to suffering (with self-kindness or reduced self-judgment), *cognitively* understand their suffering (with common humanity or reduced isolation), and *pay attention* to their suffering (with mindfulness or reduced overidentification). Thus, within the dimensions of CS and RUS the three components are not thought to be identical.

Given that support was found for use of a total score and also six separate subscale scores, the question arises—when is use of a total score versus subscale scores warranted? The nomological network observed between the six subscales and important aspects of functioning indicates that there are areas of overlap but also difference between the subscales. For instance, Körner et al. (2015) found that it was mainly isolation that predicted depression, while Alda et al. (2016) found that common humanity had the strongest association with telomere length. Moreover, evidence from neuroimaging studies suggests the various components of self-compassion have distinct brain signatures. Longe et al. (2010) found that self-critical thinking (similar to self-judgment) and self-reassurance (similar to self-kindness) were associated with different regions of brain activity. Self-criticism was associated with activity in lateral prefrontal cortex (PFC) regions and dorsal

Table 5
Goodness-of-Fit Indices for the Total and Individual Samples: Correlated Two-Bifactor Models

Country	Correlated two-bifactor CFA					Correlated two-bifactor ESEM				
	CFI	TLI	RMSEA	90% CI	WRMR	CFI	TLI	RMSEA	90% CI	WRMR
Total	.96 ^a	.95	.06	[.06-.06]	4.49	.99	.99	.04	[.03-.04]	1.20
AUS	.96	.95	.08	[.07-.08]	1.09	.99	.98	.04	[.03-.05]	.36
BRA	.97 ^a	.96	.06	[.06-.07]	1.02	.99	.98	.04	[.03-.05]	.41
CAN			No identification			.98	.97	.05	[.04-.06]	.44
CHI			No identification			.99 ^a	.99	.05	[.04-.06]	.36
FRA	.95 ^a	.94	.08	[.07-.08]	2.06	.99	.98	.04	[.04-.05]	.59
GER	.90	.88	.08	[.08-.09]	1.37	.99 ^a	.97	.04	[.03-.05]	.42
GRE			No identification			.99	.99	.04	[.03-.04]	.46
IRA			No identification			.98	.96	.04	[.03-.05]	.51
ITA	.94	.93	.08	[.07-.08]	1.31	.99	.98	.04	[.03-.05]	.40
JAP			No identification			.99 ^a	.97	.04	[.04-.05]	.55
KOR			No identification			.99	.97	.05	[.04-.06]	.40
NOR	.93	.91	.08	[.08-.09]	1.31	.99	.98	.04	[.03-.05]	.40
POR 1	.96	.96	.07	[.07-.08]	1.88	.99	.98	.05	[.04-.05]	.54
POR 2			No identification			.98	.96	.06	[.05-.06]	.44
SPA						.99	.97	.04	[.03-.05]	.44
UK 1	.94	.93	.08	[.07-.08]	1.74	.99	.99	.03	[.03-.04]	.48
UK 2	.93 ^a	.92	.07	[.06-.08]	1.29	.99	.97	.04	[.03-.05]	.44
US 1	.96	.95	.07	[.07-.08]	1.56	.99	.99	.03	[.02-.04]	.42
US 2	.91	.90	.08	[.08-.08]	1.83	.99	.97	.04	[.04-.05]	.52
US 3			No identification			.99	.99	.04	[.02-.05]	.33

Note. CFA = confirmatory factor analysis; ESEM = exploratory structural equation modeling; CFI = comparative fit index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; WRMR = weighted root mean square residual; CI = confidence interval; AUS = Australia; BRA = Brazil; CAN = Canada; CHI = China; FRA = France; GER = Germany; GRE = Greece; IRA = Iran; ITA = Italy; JAP = Japan; KOR = South Korea; NOR = Norway; POR = Portugal; SPA = Spain; UK = United Kingdom; US = United States.

^a These solutions had model identification issues, suggesting overparameterization.

anterior cingulate (dACC), linked to error processing and resolution, and also behavioral inhibition. Self-reassurance was associated with left temporal pole and insula activation, related to empathy. Mindfulness, on the other hand is linked to increased neural activation in the prefrontal cortex (PFC) and dorsal anterior cingulate cortex (dACC), associated with attentional control and emotion regulation (Young et al., 2017). These results suggest that the six components of self-compassion are not one unitary thing, nor are they two unitary things, but are six distinct but interrelated things.

Use of the subscales may have relevance for understanding the mechanisms by which self-compassion engenders well-being. Neff, Long, et al. (in press) recently explored the link of self-compassion and its components to psychological functioning in seven domains—psychopathology, positive psychological health, emotional intelligence, self-concept, body image, motivation, and interpersonal functioning. When examining the zero-order correlations between observed subscale scores and outcomes, they found that reduced self-judgment, isolation, and overidentification tended to have a stronger link to negative emotionality and self-evaluation than self-kindness, common humanity, and mindfulness, while the latter tended to have a stronger association with outcomes like emotional awareness, goal reengagement, compassion for others, and perspective-taking. For many aspects of psychological functioning, however, such as happiness, wisdom, contingent self-esteem based on approval, body appreciation, or grit, all six subscales appeared to make an equal contribution to well-being. They interpreted findings to mean that although different elements of self-compassion may differentially explain its link

with wellbeing, all are essential to the construct of self-compassion as a whole.

For most researchers, use of the SCS as a total score will be most appropriate given that self-compassion operates as a system. This view is supported by findings from intervention research indicating that self-compassion training changes all six components at the same time. The vast majority of intervention studies using a wide variety of methodologies that examined change in self-compassion have documented a simultaneous change in all six subscales of roughly the same magnitude: for example, self-compassion meditation training (e.g., Toole & Craighead, 2016); online psycho-education (e.g., Krieger, Martig, van den Brink, & Berger, 2016); Compassion Focused Therapy (e.g., Beaumont, Irons, Rayner, & Dagnall, 2016); Compassionate Mind Training (e.g., Arimitsu, 2016); and the Mindful Self-Compassion program (e.g., Neff, 2016a). Not only do self-compassion interventions impact CS and RUS to the same degree, changes in both impact outcomes similarly. Krieger, Berger, and Holtforth (2016) used cross-lagged analyses to explore whether changes in self-compassion over the course of cognitive-behavioral psychotherapy led to changes in depression, and findings were the same whether a total score or two scores representing compassionate or uncompassionate responding were examined. They interpreted findings as evidence that self-compassion should be considered an overall construct rather than two separate constructs. Similarly, Neff (2016a) found that changes in SCS subscales representing CS and RUS after 8 weeks of self-compassion training tended to be equally predictive of changes in happiness, life satisfaction, anxiety, depression, and stress.

Table 6
Standardized Factor Loadings for the Six-Factor Correlated CFA and ESEM Solutions of the Self-Compassion Scale for the Total Sample

Items	CFA		ESEM				
	SF (λ) ^a	SK (λ)	SJ (λ)	CH (λ)	IS (λ)	MI (λ)	OI (λ)
Self-kindness							
sk5	.74	.69	.04	.12	.01	.02	.01
sk12	.82	.84	.03	.02	.01	.01	.04
sk19	.84	.80	.03	.00	.00	.08	.01
sk23	.75	.26	.42	.08	.05	.31	.16
sk26	.79	.36	.34	.11	.00	.34	.18
Self-judgment							
sj1	.76	.09	.61	.05	.06	.05	.18
sj8	.74	.25	.43	.00	.13	.16	.22
sj11	.74	.09	.51	.02	.13	.05	.13
sj16	.83	.05	.49	.02	.23	.03	.20
sj21	.73	.33	.33	.01	.14	.14	.21
Common humanity							
ch3	.70	.04	.11	.45	.15	.19	.01
ch7	.65	.08	.06	.97	.04	.15	.04
ch10	.73	.00	.02	.87	.08	.06	.03
ch15	.84	.09	.05	.43	.07	.29	.06
Isolation							
is4	.79	.01	.28	.08	.43	.01	.13
is13	.79	.01	.10	.03	.97	.02	.06
is18	.72	.04	.10	.02	.90	.00	.03
is25	.79	.00	.20	.13	.37	.04	.26
Mindfulness							
mi9	.66	.13	.14	.09	.08	.53	.33
mi14	.79	.16	.15	.12	.09	.58	.17
mi17	.77	.14	.01	.16	.10	.49	.05
mi22	.72	.38	.02	.13	.08	.29	.06
Overidentification							
oi2	.82	.02	.34	.05	.20	.04	.38
oi6	.78	.03	.40	.08	.16	.01	.31
oi20	.68	.06	.07	.01	.01	.20	.69
oi24	.69	.05	.03	.02	.14	.21	.58

Note. CFA = confirmatory factor analysis; ESEM = exploratory structural equation modeling; SF = specific factor; SK = self-kindness; SJ = self-judgment; CH = common humanity; IS = isolation; MI = mindfulness; OI = overidentification; λ = standardized factor loadings. Target factor loadings are in bold. Nonsignificant parameters ($p \geq .05$) are italicized.

^a Each item loaded on its respective specific factor, while cross-loadings were constrained to zero.

These findings suggest that self-compassion is experienced holistically. They also buttress current study findings supporting the use of a total SCS score to represent self-compassion as defined by Neff (2003b). Perhaps most importantly, they highlight why there is so much excitement about the construct of self-compassion in the field of psychology: It is a skill that can be learned (Neff & Germer, 2013). For researchers who are primarily interested in self-compassion as a trainable mind-state, therefore, use of a total score is probably most appropriate. For those more interested in unpacking the mechanisms of how self-compassion enhances well-being, however, it may be useful to examine the six constituent components themselves.

An important contribution of the present investigation is the finding that self-compassion is better represented with a single continuum rather than two distinct dimensions of CS and RUS. This notion was supported by the fact that the positively and negatively valenced items loaded on the general factor in a similar magnitude in the model including one general factor, whereas these loadings were weak in the model with two correlated general

factors. It should be noted that the separation of positive and negative items sometimes results from a clustering effect where items with a similar valence load onto separate factors, basically forming method factors that mostly originate from the positive versus negative wording of the items (Crego & Widiger, 2014). This has been shown in research on self-esteem (Greenberger, Chen, Dmitrieva, & Farruggia, 2003; Marsh, 1996), for instance, where method factors emerged as a result of item wording. Generally, wording effects would be interpreted as substantively irrelevant artifacts, but in the case of the SCS, we do not believe that the separation of positively and negatively valenced items is a result of item wording only. Rather, the distinction between compassionate and reduced uncompassionate responding toward oneself is conceptually meaningful and substantially contributes to the global self-compassion factor. Self-compassion can be conceptualized as a holistic state of mind representing the balance of CS and RUS along the three basic dimensions of emotional responding, cognitive understanding, and paying attention to personal distress.

Table 7
Standardized Factor Loadings for the Bifactor CFA and ESEM Solutions of the Self-Compassion Scale for the Total Sample

Items	Bifactor-CFA		Bifactor-ESEM						
	GF (λ)	SF (λ) ^a	GF (λ)	SK (λ)	SJ (λ)	CH (λ)	IS (λ)	MI (λ)	OI (λ)
Self-kindness									
sk5	.59	.50	.58	.47	.04	.17	.05	.12	.06
sk12	.66	.54	.64	.56	.01	.11	.03	.12	.03
sk19	.67	.53	.68	.50	.03	.08	.07	.12	.07
sk23	.66	.27	.72	.06	.04	.01	.15	.08	.24
sk26	.67	.33	.73	.13	.13	.06	.19	.12	.26
Self-judgment									
sj1	.65	.42	.67	.06	.44	.12	.02	.12	.05
sj8	.62	.44	.66	.04	.20	.13	.06	.25	.13
sj11	.64	.39	.70	.09	.15	.12	.01	.14	.03
sj16	.72	.39	.75	.10	.23	.13	.11	.12	.09
sj21	.63	.36	.67	.11	.07	.10	.05	.25	.13
Common humanity									
ch3	.51	.39	.46	.09	.11	.38	.05	.24	.03
ch7	.38	.72	.36	.08	.07	.73	.05	.04	.02
ch10	.48	.64	.44	.11	.03	.65	.07	.10	.01
ch15	.63	.36	.58	.08	.07	.35	.05	.27	.12
Isolation									
is4	.69	.26	.66	.08	.20	.05	.26	.06	.10
is13	.64	.58	.64	.06	.02	.06	.58	.00	.05
is18	.56	.57	.57	.07	.03	.06	.55	.02	.07
is25	.69	.25	.67	.07	.11	.00	.24	.10	.20
Mindfulness									
mi9	.53	.45	.50	.09	.15	.15	.07	.43	.16
mi14	.65	.55	.59	.12	.10	.19	.00	.52	.04
mi17	.64	.39	.61	.08	.07	.17	.03	.40	.06
mi22	.61	.27	.55	.25	.07	.17	.04	.28	.12
Overidentification									
oi2	.75	.23	.69	.05	.34	.07	.17	.04	.27
oi6	.71	.17	.68	.07	.25	.07	.11	.08	.19
oi20	.58	.57	.59	.12	.06	.05	.05	.03	.50
oi24	.60	.42	.60	.11	.08	.03	.12	.05	.41

Note. CFA = confirmatory factor analysis; ESEM = exploratory structural equation modeling; GF = general factor of self-compassion; SF = specific factor; ¹ = Each item loaded on its respective specific factor, while cross-loadings were constrained to zero; SK = self-kindness; SJ = self-judgment; CH = common humanity; IS = isolation; MI = mindfulness; OI = overidentification; λ = standardized factor loadings. Target factor loadings are in bold. Nonsignificant parameters ($p \geq .05$) are italicized.

Limitations and Future Directions

While this is one of the most comprehensive examinations of the factor structure of the SCS conducted to date, there were some limitations. For instance, the populations included were majority female and mainly community and student samples; only one meditator and two clinical samples were included. Fit in these latter samples was excellent, providing some confidence in use of the SCS with these populations. Still, it will be important to make sure the factor structure replicates in specific types of populations (anxious, eating disordered, etc.). Also, although findings support use of the SCS in different cultures, reliability coefficients, and model fit did vary somewhat across samples (less so for our chosen models). Also, in some countries (e.g., China and Japan) multiple measurement models presented identification issues, and it should be investigated whether these issues relate to model misspecification or sampling-specific errors. Potential differences in the SCS structure should also be addressed with analyses of invariance across culture, population type, age, and sex, as these may be addi-

tional sources of meaningful variation in the SCS that should be understood. (These analyses are being conducted for the current dataset and will be presented in a separate article; Neff et al., 2018)

Given the superiority demonstrated by the ESEM models over CFA models, results suggest that future attempts to validate translations of the SCS or to examine the properties of the SCS in specific populations should use this approach (syntax files are available for interested readers in Appendix 5 of the supplementary materials). Additional studies are also needed to examine the criterion-validity of test score interpretations using this improved representation to better capture the meaning of the subscales once the global level of self-compassion is accounted for. Use of the bifactor ESEM framework aligns with the proposition of Marsh and Hau (2007) who emphasized the need for the use of latent variable models that, compared with observed variables, more accurately define constructs with the explicit inclusion of measurement errors related to the imperfect items. Bifactor ESEM models are rather complex and

Table 8
Standardized Factor Loadings for the Two-Bifactor CFA and ESEM Solutions of the Self-Compassion Scale for the Total Sample

Items	Two-bifactor-CFA			Two-bifactor-ESEM							
	CS (λ)	RUS (λ)	SF (λ) ¹	CS (λ)	RUS (λ)	SK (λ)	SJ (λ)	CH (λ)	IS (λ)	MI (λ)	OI (λ)
Self-kindness											
sk5	.70		.32	.43		.24	.31	.34	.10	.37	.06
sk12	.78		.39	.48		.32	.39	.30	.13	.40	.09
sk19	.80		.31	.46		.25	.42	.28	.11	.44	.07
sk23	.77		.21	.31		.22	.49	.26	.14	.40	.05
sk26	.79		.08	.36		.17	.45	.32	.09	.47	.03
Self-judgment											
sj1		.72	.28		.16	.00	.71	.07	.30	.06	.26
sj8		.69	.33		.15	.11	.67	.06	.21	.07	.23
sj11		.70	.24		.09	.08	.64	.09	.22	.17	.22
sj16		.79	.22		.06	.04	.66	.08	.34	.15	.31
sj21		.69	.23		.35	.18	.67	.09	.09	.14	.19
Common humanity											
ch3	.56		.31	.07		.06	.09	.48	.18	.39	.09
ch7	.44		.70	.03		.07	.04	.79	.04	.16	.06
ch10	.54		.59	.08		.07	.10	.75	.05	.23	.11
ch15	.68		.27	.09		.00	.22	.50	.13	.49	.05
Isolation											
is4		.74	.14		.10	.01	.50	.11	.45	.13	.28
is13		.69	.52		.32	.03	.31	.09	.68	.20	.21
is18		.62	.52		.32	.02	.26	.07	.63	.15	.21
is25		.74	.12		.18	.02	.46	.17	.39	.11	.37
Mindfulness											
mi9	.59		.36	.09		.06	.08	.26	.07	.58	.29
mi14	.71		.48	.01		.13	.15	.29	.16	.76	.15
mi17	.70		.26	.12		.01	.21	.32	.17	.60	.13
mi22	.67		.12	.27		.07	.22	.32	.14	.49	.04
Overidentification											
oi2		.80	.05		.07	.10	.52	.08	.41	.15	.47
oi6		.76	.02		.04	.02	.56	.11	.31	.11	.37
oi20		.63	.86		.19	.01	.28	.08	.16	.20	.68
oi24		.66	.21		.21	.01	.28	.11	.23	.24	.57

Note. CFA = confirmatory factor analysis; ESEM = exploratory structural equation modeling; CS = general factor representing Compassionate Self-Responding; RUS = general factor representing Reduced Uncompassionate Self-responding; SF = specific factor; ¹ = Each item loaded on its respective specific factor, while cross-loadings were constrained to zero; SK = self-kindness; SJ = self-judgment; CH = common humanity; IS = isolation; MI = mindfulness; OI = overidentification; λ = standardized factor loadings. Target factor loadings are in bold. Nonsignificant parameters ($p \geq .05$) are italicized.

sometimes difficult to incorporate into predictive models because of the relatively high number of estimated parameters, but one can construct separate measurement models and “translate” these measurement models into factor scores saved from these preliminary measurement models that are better at preserving the a priori nature of the constructs compared with observed variables (Gillet, Morin, Cougot, & Gagné, 2017; Morin, Meyer, Creusier, & Biétry, 2016). Neff, Tóth-Király, and Colosimo (2018) successfully used this approach to examine the incremental validity of self-compassion and neuroticism in predicting wellbeing. (Syntax for saving factor scores is also included in Appendix 5 of the supplementary materials.)

Note that while the current study was aimed at examining the validity of test score interpretations on the SCS as a measure of Neff’s (2003b) conceptualization of self-compassion, in no way can it speak to the issue of whether this definition or measurement of self-compassion is superior to others. For example, Social Mentality Theory (SMT; Gilbert, 1989, 2005) posits that self-compassion is a state of mind that emerges from mammalian bio-social roles involving caregiving and care-seeking,

while self-criticism emerges from evolved social roles that protect us from social threats. To this end Gilbert and colleagues developed the Forms of Self-Criticism and Self-Reassurance Scales (Gilbert et al., 2004) to measure these two ways of relating to oneself. More recently, Gilbert and colleagues (Gilbert et al., 2017) have developed a model of compassion for self, for others, and from others, based on the broadly used definition of compassion as sensitivity to suffering with a commitment to try to alleviate it (Goetz, Keltner, & Simon-Thomas, 2010). They developed the Compassion Engagement and Action Scales, including self-compassion and other compassion scales with items tapping into engagement with distress (e.g., tolerating and being sensitive to distress) and the motivation to alleviate that distress (e.g., thinking about and taking actions to help). Notably, the scales do not include kindness/concern or shared humanity as a feature of compassion. As with the SCS (Neff & Pommier, 2013), scale scores measuring compassion for self and others are only weakly correlated, with higher levels of compassion being reported for others than the self. It is unclear if the desire to alleviate distress

Table 9
Reliability Estimates for the Bifactor ESEM Model for the Total and Individual Samples

Country	Bifactor			
	ω	ω_H	GF	SF
Total	.96	.91	.95	.05
AUS	.98	.93	.95	.05
BRA	.97	.91	.94	.06
CAN	.96	.88	.92	.08
CHI		Negative residual variance		
FRA	.97	.92	.95	.05
GER	.96	.88	.92	.08
GRE	.97	.91	.94	.06
IRA	.93	.85	.91	.08
ITA	.96	.89	.93	.07
JAP		Negative residual variance		
KOR	.95	.82	.86	.13
NOR	.96	.89	.93	.07
POR 1	.97	.90	.93	.07
POR 2	.96	.90	.94	.06
SPA	.94	.83	.88	.11
UK 1	.97	.92	.95	.05
UK 2	.96	.89	.93	.07
US 1	.97	.93	.96	.04
US 2	.95	.87	.92	.08
US 3	.98	.93	.95	.05

Note. ω = omega; ω_H = omega hierarchical; GF = Reliable variance explained by the general factor; SF = Reliable variance explained by the specific factors; AUS = Australia; BRA = Brazil; CAN = Canada; CHI = China; FRA = France; GER = Germany; GRE = Greece; IRA = Iran; ITA = Italy; JAP = Japan; KOR = South Korea; NOR = Norway; POR = Portugal; SPA = Spain; UK = United Kingdom; US = United States.

operates the same way for self and others, however, given that the desire to alleviate personal distress overlaps with resistance to distress. Resistance can exacerbate psychopathology, which is why mindfulness-based clinical approaches such as Acceptance and Commitment Therapy (Hayes, Strosahl, & Wilson, 1999) and Mindfulness-Based Cognitive Therapy (Segal, Williams, & Teasdale, 2012) are aimed at reducing resistance to personal distress.

Strauss et al. (2016) propose that measures of compassion should include five key elements: (1) Recognizing suffering; (2) Understanding the universality of suffering in human experience; (3) Feeling concern for the person suffering; (4) Tolerating uncomfortable feelings in response to suffering, so remaining open to and accepting of the person suffering; and (5) Motivation to alleviate suffering. While the SCS taps into most of these elements, no items explicitly address the motivation to alleviate suffering. This is because the motivation to alleviate the self's suffering is easily conflated with resistance to personal distress (undermining the fourth element) in a way that is less problematic in measures of compassion for others. Still, future research might fruitfully explore whether adding items to the SCS that are focused on the motivation to help and support oneself in times of distress could strengthen the measurement of self-compassion.

To summarize, in the 20 diverse samples we examined, the excellent fit of single-bifactor ESEM and six-factor correlated ESEM models found across samples strongly supports the conclusion that self-compassion as measured by the SCS can be

viewed as a general construct (explaining 95% of the reliable variance in item responding), comprised of six separate components. While the constituent elements of self-compassion are distinct and can be measured separately, they operate in tandem, as suggested by the large body of research examining self-compassion interventions. Hopefully these findings can help put some of the controversy over the factor structure of the SCS to rest: A total score rather than two separate scores should be used.

References

- Alda, M., Puebla-Guedea, M., Rodero, B., Demarzo, M., Montero-Marín, J., Roca, M., & Garcia-Campayo, J. (2016). Zen meditation, length of telomeres, and the role of experiential avoidance and compassion. *Mindfulness*, 7, 651–659. <http://dx.doi.org/10.1007/s12671-016-0500-5>
- Arimitsu, K. (2014). Development and validation of the Japanese version of the Self-Compassion Scale. *The Japanese Journal of Psychology*, 85, 50–59. <http://dx.doi.org/10.4992/jjpsy.85.50>
- Arimitsu, K. (2016). The effects of a program to enhance self-compassion in Japanese individuals: A randomized controlled pilot study. *The Journal of Positive Psychology*, 11, 559–571. <http://dx.doi.org/10.1080/17439760.2016.1152593>
- Arimitsu, K., Aoki, Y., Furukita, M., Tada, A., & Togashi, R. (2016). Construction and validation of a short form of the Japanese version of the Self-Compassion Scale. *Komazawa Annual Reports of Psychology*, 18, 1–8.
- Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, 16, 397–438. <http://dx.doi.org/10.1080/10705510903008204>
- Asparouhov, T., Muthén, B., & Morin, A. J. S. (2015). Bayesian structural equation modeling with cross-loadings and residual covariances: Comments on Stromeier et al. *Journal of Management*, 41, 1561–1577. <http://dx.doi.org/10.1177/0149206315591075>
- Azizi, A., Mohammadkhani, P., Lotfi, S., & Bahramkhani, M. (2013). The validity and reliability of the Iranian version of the Self-Compassion Scale. *Iranian Journal of Clinical Psychology*, 2, 17–23.
- Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, 16, 74–94. <http://dx.doi.org/10.1007/BF02723327>
- Bandalos, D. L. (2014). Relative performance of categorical diagonally weighted least squares and robust maximum likelihood estimation. *Structural Equation Modeling*, 21, 102–116. <http://dx.doi.org/10.1080/10705511.2014.859510>
- Beaumont, E., Irons, C., Rayner, G., & Dagnall, N. (2016). Does Compassion-Focused Therapy training for health care educators and providers increase self-compassion and reduce self-persecution and self-criticism? *The Journal of Continuing Education in the Health Professions*, 36, 4–10. <http://dx.doi.org/10.1097/CEH.0000000000000023>
- Benda, J., & Reichová, A. (2016). Psychometrické charakteristiky České verze Self-Compassion Scale (SCS-CZ) [Psychometric characteristics of the Czech version of the Self-Compassion Scale]. *Československá Psychologie*, 60, 20–36.
- Bento, E., Xavier, S., Azevedo, J., Marques, M., Freitas, V., Soares, M. J., . . . Pereira, A. T. (2016). Validation of the self-compassion scale in a community sample of Portuguese pregnant women. *European Psychiatry*, 33, S282–S283. <http://dx.doi.org/10.1016/j.eurpsy.2016.01.598>
- Birnie, K., Speca, M., & Carlson, L. E. (2010). Exploring self-compassion and empathy in the context of Mindfulness-based Stress Reduction (MBSR). *Stress and Health*, 26, 359–371. <http://dx.doi.org/10.1002/smi.1305>

- 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*, 696–707. <http://dx.doi.org/10.1037/cou0000211>
- Browne, M. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research, 36*, 111–150. http://dx.doi.org/10.1207/S15327906MBR3601_05
- Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality, 80*, 796–846. <http://dx.doi.org/10.1111/j.1467-6494.2011.00749.x>
- Castilho, P., & Pinto-Gouveia, J. (2011). Self-compassion: Validation of the Portuguese version of the Self-Compassion Scale and its relationship to adverse childhood experiences, social comparison and psychopathology. *Psychologica, 54*, 203–291. http://dx.doi.org/10.14195/1647-8606_54_8
- Castilho, P., Pinto-Gouveia, J., & Duarte, J. (2015). Evaluating the multi-factor structure of the long and short versions of the self-compassion scale in a clinical sample. *Journal of Clinical Psychology, 71*, 856–870. <http://dx.doi.org/10.1002/jclp.22187>
- Chang, E. C., Yu, T., Jilani, Z., Fowler, E. E., Yu, E. A., Lin, J., & Hirsch, J. K. (2015). Under assault: Understanding the impact of sexual assault on the relation between hope and suicidal risk in college students. *Journal of Social and Clinical Psychology, 34*, 221–238. <http://dx.doi.org/10.1521/jscp.2015.34.3.221>
- Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research, 41*, 189–225. http://dx.doi.org/10.1207/s15327906mbr4102_5
- Chen, J., Yan, L., & Zhou, L. (2011). Reliability and validity of Chinese version of Self-Compassion Scale. *Chinese Journal of Clinical Psychology, 19*, 734–736.
- 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*, 618–628. <http://dx.doi.org/10.1007/s12671-017-0803-1>
- Coroiu, A., Kwakkenbos, L., Moran, C., Thombs, B., Albani, C., Bourkas, S., . . . Körner, A. (2018). Structural validation of the Self-Compassion Scale with a German general population sample. *PLoS ONE, 13*, e0190771. <http://dx.doi.org/10.1371/journal.pone.0190771>
- Costa, J., Marôco, J., Pinto-Gouveia, J., Ferreira, C., & Castilho, P. (2016). Validation of the psychometric properties of the Self-Compassion Scale. *Clinical Psychology & Psychotherapy, 23*, 460–468. <http://dx.doi.org/10.1002/cpp.1974>
- Crego, C., & Widiger, T. A. (2014). Psychopathy, DSM–5, and a caution. *Personality Disorders, 5*, 335–347. <http://dx.doi.org/10.1037/per0000078>
- Crocker, J., & Park, L. E. (2004). The costly pursuit of self-esteem. *Psychological Bulletin, 130*, 392–414. <http://dx.doi.org/10.1037/0033-2909.130.3.392>
- Cunha, M., Xavier, A., & Castilho, P. (2016). Understanding self-compassion in adolescents: Validation study of the Self-Compassion Scale. *Personality and Individual Differences, 93*, 56–62. <http://dx.doi.org/10.1016/j.paid.2015.09.023>
- Deniz, M., Kesici, Ş., & Sümer, A. S. (2008). The validity and reliability of the Turkish version of the Self-Compassion Scale. *Social Behavior and Personality, 36*, 1151–1160. <http://dx.doi.org/10.2224/sbp.2008.36.9.1151>
- de Souza, L. K., Ávila-Souza, J., & Gauer, G. (2016). Escala de Autocompaixão [Self-compassion Scale]. In C. S. Hutz (Ed.), *Avaliação em Psicologia Positiva* (pp. 169–177). São Paulo: Hogrefe/CETEP.
- de Souza, L. K., & Hutz, C. S. (2016). Adaptation of the Self-Compassion Scale for use in Brazil: Evidences of construct validity. *Trends in Psychology, 24*, 159–172.
- Dundas, I., Svendsen, J. L., Wiker, A. S., Granli, K. V., & Schanche, E. (2016). Self-compassion and depressive symptoms in a Norwegian student sample. *Nordic Psychology, 68*, 58–72. <http://dx.doi.org/10.1080/19012276.2015.1071203>
- Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G. R. Hancock & R. D. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 269–314). Charlotte, NC: Information Age Publishing.
- Furr, R. M., & Bacharach, V. R. (2008). *Psychometrics: An introduction*. Thousand Oaks, CA: Sage.
- García-Campayo, J., Navarro-Gil, M., Andrés, E., Montero-Marin, J., López-Artal, L., & Demarzo, M. M. (2014). Validation of the Spanish versions of the long (26 items) and short (12 items) forms of the Self-Compassion Scale (SCS). *Health and Quality of Life Outcomes, 12*, 4. <http://dx.doi.org/10.1186/1477-7525-12-4>
- Ghorbani, N., Chen, Z., Saeedi, Z., Behjati, Z., & Watson, P. J. (2013). Saktare Ameli Meghyase Shafeghat e Khod dar Iran [Factorial structure of Self-Compassion Scale in Iran]. *Pazhohesh Haye KARBORDI Ravanshenakhti, 4*, 29–41.
- Gignac, G. E. (2016). The higher-order model imposes a proportionality constraint: That is why the bifactor model tends to fit better. *Intelligence, 55*, 57–68. <http://dx.doi.org/10.1016/j.intell.2016.01.006>
- Gilbert, P. (1989). *Human nature and suffering*. London, UK: Routledge.
- Gilbert, P. (2005). *Compassion: Conceptualizations, research and use in psychotherapy*. London, UK: Brunner-Routledge.
- Gilbert, P., Catarino, F., Duarte, C., Matos, M., Kolts, R., Stubbs, J., . . . Basran, J. (2017). The development of compassionate engagement and action scales for self and others. *Journal of Compassionate Health Care, 4*, 4. <http://dx.doi.org/10.1186/s40639-017-0033-3>
- Gilbert, P., Clarke, M., Hempel, S., Miles, J. N., & Irons, C. (2004). Criticizing and reassuring oneself: An exploration of forms, styles and reasons in female students. *British Journal of Clinical Psychology, 43*, 31–50. <http://dx.doi.org/10.1348/01446650472812959>
- 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*, 239–255. <http://dx.doi.org/10.1348/147608310X526511>
- Gillet, N., Morin, A. J. S., Cougot, B., & Gagné, M. (2017). Workaholism profiles: Associations with determinants, correlates, and outcomes. *Journal of Occupational and Organizational Psychology, 90*, 559–586. <http://dx.doi.org/10.1111/joop.12185>
- Goetz, J. L., Keltner, D., & Simon-Thomas, E. (2010). Compassion: An evolutionary analysis and empirical review. *Psychological Bulletin, 136*, 351–374. <http://dx.doi.org/10.1037/a0018807>
- Greenberger, E., Chen, C., Dmitrieva, J., & Farruggia, S. P. (2003). Item-wording and the dimensionality of the Rosenberg Self-Esteem Scale: Do they matter? *Personality and Individual Differences, 35*, 1241–1254. [http://dx.doi.org/10.1016/S0191-8869\(02\)00331-8](http://dx.doi.org/10.1016/S0191-8869(02)00331-8)
- Halamová, J., Kanovský, M., & Pacúchová, M. (2018). Self-Compassion Scale: IRT psychometric analysis, validation, and factor structure–Slovak translation. *Psychologica Belgica, 57*, 190–209. <http://dx.doi.org/10.5334/pb.398>
- Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and commitment therapy* (p. 6). New York, NY: Guilford Press.
- Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424–453. <http://dx.doi.org/10.1037/1082-989X.3.4.424>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. <http://dx.doi.org/10.1080/10705519909540118>
- Hupfeld, J., & Ruffieux, N. (2011). Validation of a German version of the self-compassion scale (SCS-D). *Zeitschrift für Klinische Psychologie*

- und *Psychotherapie*, 40, 115–123. <http://dx.doi.org/10.1026/1616-3443/a000088>
- Karakasidou, E., Pezirkianidis, C., Galanakis, M., & Stalikas, A. (2017). Validity, reliability and factorial structure of the Self Compassion Scale in the Greek population. *Journal of Psychology & Psychotherapy*, 7, 313.
- Kelly, A. C., Wisniewski, L., Martin-Wagar, C., & Hoffman, E. (2017). Group-based Compassion-Focused Therapy as an adjunct to outpatient treatment for eating disorders: A pilot randomized controlled trial. *Clinical Psychology & Psychotherapy*, 24, 475–487. <http://dx.doi.org/10.1002/cpp.2018>
- Kim, K. E., Yi, G. D., Cho, Y. R., Chai, S. H., & Lee, W. K. (2008). The validation study of the Korean version of the self-compassion scale. *Han'gug Simlihag Hoeji Geon'gang*, 13, 1023–1044. <http://dx.doi.org/10.17315/kjhp.2008.13.4.012>
- Körner, A., Coroiu, A., Copeland, L., Gomez-Garibello, C., Albani, C., Zenger, M., & Brähler, E. (2015). The role of self-compassion in buffering symptoms of depression in the general population. *PLoS ONE*, 10, e0136598. <http://dx.doi.org/10.1371/journal.pone.0136598>
- Kotsou, I., & Leys, C. (2016). Self-Compassion Scale (SCS): Psychometric properties of the French translation and its relations with psychological well-being, affect and depression. *PLoS ONE*, 11, e0152880. <http://dx.doi.org/10.1371/journal.pone.0152880>
- Krieger, T., Berger, T., & Holtforth, M. G. (2016). The relationship of self-compassion and depression: Cross-lagged panel analyses in depressed patients after outpatient therapy. *Journal of Affective Disorders*, 202, 39–45. <http://dx.doi.org/10.1016/j.jad.2016.05.032>
- Krieger, T., Hermann, H., Zimmermann, J., & Grosse Holtforth, M. (2015). Associations of self-compassion and global self-esteem with positive and negative affect and stress reactivity in daily life: Findings from a smart phone study. *Personality and Individual Differences*, 87, 288–292. <http://dx.doi.org/10.1016/j.paid.2015.08.009>
- Krieger, T., Martig, D. S., van den Brink, E., & Berger, T. (2016). Working on self-compassion online: A proof of concept and feasibility study. *Internet Interventions*, 6, 64–70. <http://dx.doi.org/10.1016/j.invent.2016.10.001>
- Kyeong, L. W. (2013). Self-compassion as a moderator of the relationship between academic burn-out and psychological health in Korean cyber university students. *Personality and Individual Differences*, 54, 899–902. <http://dx.doi.org/10.1016/j.paid.2013.01.001>
- Lee, W. K., & Lee, K. (2010). The validation study of the Korean version of the Self-Compassion Scale with adult women in the community. *Journal of Korean Neuropsychiatric Association*, 49, 193–200.
- Longe, O., Maratos, F. A., Gilbert, P., Evans, G., Volker, F., Rockliff, H., & Rippon, G. (2010). Having a word with yourself: Neural correlates of self-criticism and self-reassurance. *NeuroImage*, 49, 1849–1856. <http://dx.doi.org/10.1016/j.neuroimage.2009.09.019>
- López, A., Sanderman, R., Smink, A., Zhang, Y., van Sonderen, E., Ranchor, A., & Schroevers, M. J. (2015). A reconsideration of the Self-Compassion Scale's total score: Self-compassion versus self-criticism. *PLoS ONE*, 10, e0132940. <http://dx.doi.org/10.1371/journal.pone.0132940>
- Mantzios, M., Wilson, J. C., & Giannou, K. (2015). Psychometric properties of the Greek versions of the self-compassion and mindful attention and awareness scales. *Mindfulness*, 6, 123–132. <http://dx.doi.org/10.1007/s12671-013-0237-3>
- Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifacts? *Journal of Personality and Social Psychology*, 70, 810–819. <http://dx.doi.org/10.1037/0022-3514.70.4.810>
- Marsh, H. W., Hau, K.-T., & Grayson, D. (2005). Goodness of fit evaluation. In A. Maydeu-Olivares & J. McArdle (Eds.), *Contemporary psychometrics* (pp. 275–340). Mahwah, NJ: Erlbaum.
- 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, 151–170. <http://dx.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*, 11, 320–341. http://dx.doi.org/10.1207/s15328007sem1103_2
- Marsh, H. W., Liem, G. A. D., Martin, A. J., Morin, A. J. S., & Nagengast, B. (2011). Methodological measurement fruitfulness of exploratory structural equation modeling (ESEM): New approaches to key substantive issues in motivation and engagement. *Journal of Psychoeducational Assessment*, 29, 322–346. <http://dx.doi.org/10.1177/0734282911406657>
- 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, 85–110. <http://dx.doi.org/10.1146/annurev-clinpsy-032813-153700>
- Montero-Marín, J., Gaete, J., Demarzo, M., Rodero, B., Lopez, L. C. S., & García-Campayo, J. (2016). Self-criticism: A measure of uncompassionate behaviors toward the self, based on the negative components of the self-compassion scale. *Frontiers in Psychology*, 7, 1281. <http://dx.doi.org/10.3389/fpsyg.2016.01281>
- Morin, A. J. S., Arens, A. K., & Marsh, H. W. (2016a). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling*, 23, 116–139. <http://dx.doi.org/10.1080/10705511.2014.961800>
- Morin, A. J. S., Arens, A. K., Tran, A., & Caci, H. (2016b). Exploring sources of construct-relevant multidimensionality in psychiatric measurement: A tutorial and illustration using the Composite Scale of Morningness. *International Journal of Methods in Psychiatric Research*, 25, 277–288. <http://dx.doi.org/10.1002/mpr.1485>
- 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, 438–454. <http://dx.doi.org/10.1080/10705511.2015.1116077>
- Morin, A. J. S., & Maïano, C. (2011). Cross-validation of the short form of the physical self-inventory (PSI-S) using exploratory structural equation modeling (ESEM). *Psychology of Sport and Exercise*, 12, 540–554. <http://dx.doi.org/10.1016/j.psychsport.2011.04.003>
- Morin, A. J. S., Marsh, H. W., & Nagengast, B. (2013). Exploratory structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 395–436). Charlotte, NC: Information Age Publishing, Inc.
- Morin, A. J. S., Meyer, J. P., Creusier, J., & Biétry, F. (2016). Multiple-group analysis of similarity in latent profile solutions. *Organizational Research Methods*, 19, 231–254. <http://dx.doi.org/10.1177/1094428115621148>
- Morin, A. J. S., Myers, N. D., & Lee, S. (in press). Modern factor analytic techniques: Bifactor models, exploratory structural equation modeling (ESEM) and bifactor-ESEM. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (4th ed.). New York, NY: Wiley.
- 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, 373–383. <http://dx.doi.org/10.1002/cpp.2005>
- Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Author.

- Neff, K. D. (2003a). Development and validation of a scale to measure self-compassion. *Self and Identity*, 2, 223–250. <http://dx.doi.org/10.1080/15298860309027>
- Neff, K. D. (2003b). Self-compassion: An alternative conceptualization of a healthy attitude toward oneself. *Self and Identity*, 2, 85–101. <http://dx.doi.org/10.1080/15298860309032>
- Neff, K. D. (2011). Self-compassion, self-esteem, and well-being. *Social and Personality Psychology Compass*, 5, 1–12. <http://dx.doi.org/10.1111/j.1751-9004.2010.00330.x>
- Neff, K. D. (2016a). The Self-Compassion Scale is a valid and theoretically coherent measure of self-compassion. *Mindfulness*, 7, 264–274. <http://dx.doi.org/10.1007/s12671-015-0479-3>
- Neff, K. D. (2016b). Does self-compassion entail reduced self-judgment, isolation, and over-identification? *Mindfulness*, 7, 791–797. <http://dx.doi.org/10.1007/s12671-016-0531-y>
- Neff, K. D., & Germer, C. K. (2013). A pilot study and randomized controlled trial of the mindful self-compassion program. *Journal of Clinical Psychology*, 69, 28–44. <http://dx.doi.org/10.1002/jclp.21923>
- Neff, K. D., & Germer, C. (2017). Self-compassion and psychological wellbeing. In J. Doty (Ed.), *Oxford handbook of compassion science* (pp. 371–386). New York, NY: Oxford University Press.
- Neff, K. D., Kirkpatrick, K., & Rude, S. S. (2007). Self-compassion and adaptive psychological functioning. *Journal of Research in Personality*, 41, 139–154. <http://dx.doi.org/10.1016/j.jrp.2006.03.004>
- Neff, K. D., Long, P., Knox, M. C., Davidson, O., Kuchar, A., Costigan, A., . . . Breines, J. (in press). The forest and the trees: Examining the association of self-compassion and its positive and negative components with psychological functioning. *Self and Identity*.
- Neff, K. D., & Pommier, E. (2013). The relationship between self-compassion and other-focused concern among college undergraduates, community adults, and practicing meditators. *Self and Identity*, 12, 160–176. <http://dx.doi.org/10.1080/15298868.2011.649546>
- Neff, K. D., Tóth-Király, I., & Colosimo, K. (2018). Self-compassion is best measured as a global construct and is overlapping with but distinct from neuroticism: A response to Pfattheicher, Geiger, Hartung, Weiss, and Schindler (2017). *European Journal of Personality*. Advance online publication. <http://dx.doi.org/10.1002/per.2148>
- Neff, K. D., Tóth-Király, I., Yarnell, L. M., Arimitsu, K., Castilho, P., Ghorbani, N., . . . Mantzios, M. (2018). *Examining invariance in the factor structure of the Self-Compassion Scale in 20 international samples*. Manuscript in preparation.
- Neff, K. D., & Vonk, R. (2009). Self-compassion versus global self-esteem: Two different ways of relating to oneself. *Journal of Personality*, 77, 23–50. <http://dx.doi.org/10.1111/j.1467-6494.2008.00537.x>
- Neff, K. D., Whittaker, T., & Karl, A. (2017). Evaluating the factor structure of the Self-Compassion Scale in four distinct populations: Is the use of a total self-compassion score justified? *Journal of Personality Assessment*, 99, 596–607. <http://dx.doi.org/10.1080/00223891.2016.1269334>
- Petrocchi, N., Ottaviani, C., & Couyoumdjian, A. (2013). Dimensionality of self-compassion: Translation and construct validation of the self-compassion scale in an Italian sample. *Journal of Mental Health*, 23, 72–77.
- Pfattheicher, S., Geiger, M., Hartung, J., Weiss, S., & Schindler, S. (2017). Old wine in new bottles? The case of self-compassion and neuroticism. *European Journal of Personality*, 31, 160–169. <http://dx.doi.org/10.1002/per.2097>
- Porges, S. W. (2001). The polyvagal theory: Phylogenetic substrates of a social nervous system. *International Journal of Psychophysiology*, 42, 123–146. [http://dx.doi.org/10.1016/S0167-8760\(01\)00162-3](http://dx.doi.org/10.1016/S0167-8760(01)00162-3)
- Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, 21, 173–184. <http://dx.doi.org/10.1177/01466216970212006>
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47, 667–696. <http://dx.doi.org/10.1080/00273171.2012.715555>
- Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, 95, 129–140. <http://dx.doi.org/10.1080/00223891.2012.725437>
- Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, 92, 544–559. <http://dx.doi.org/10.1080/00223891.2010.496477>
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016a). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, 98, 223–237. <http://dx.doi.org/10.1080/00223891.2015.1089249>
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016b). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, 21, 137–150. <http://dx.doi.org/10.1037/met0000045>
- Segal, Z. V., Williams, J. M. G., & Teasdale, J. D. (2012). *Mindfulness-based cognitive therapy for depression*. New York, NY: Guilford Press.
- Sirois, F. M., Kitner, R., & Hirsch, J. K. (2015). Self-compassion, affect, and health-promoting behaviors. *Health Psychology*, 34, 661–669. <http://dx.doi.org/10.1037/hea0000158>
- Strauss, C., Lever Taylor, B., Gu, J., Kuyken, W., Baer, R., Jones, F., & Cavanagh, K. (2016). What is compassion and how can we measure it? A review of definitions and measures. *Clinical Psychology Review*, 47, 15–27. <http://dx.doi.org/10.1016/j.cpr.2016.05.004>
- Stutts, L., Leary, M., Zeveney, A., & Hufnagle, A. (in press). A longitudinal analysis of the relationship between self-compassion and the psychological effects of perceived stress. *Self and Identity*.
- Svensden, J. L., Osnes, B., Binder, P. E., Dundas, I., Visted, E., Nordby, H., . . . Sørensen, L. (2016). Trait self-compassion reflects emotional flexibility through an association with high vagally mediated heart rate variability. *Mindfulness*, 7, 1103–1113. <http://dx.doi.org/10.1007/s12671-016-0549-1>
- Toole, A. M., & Craighead, L. W. (2016). Brief self-compassion meditation training for body image distress in young adult women. *Body Image*, 19, 104–112. <http://dx.doi.org/10.1016/j.bodyim.2016.09.001>
- Tóth-Király, I., Bőthe, B., & Orosz, G. (2017). Exploratory structural equation modeling analysis of the Self-Compassion Scale. *Mindfulness*, 8, 881–892. <http://dx.doi.org/10.1007/s12671-016-0662-1>
- Tóth-Király, I., Bőthe, B., Rigó, A., & Orosz, G. (2017). An illustration of the exploratory structural equation modeling (ESEM) framework on the Passion Scale. *Frontiers in Psychology*, 8, 1968. <http://dx.doi.org/10.3389/fpsyg.2017.01968>
- Tóth-Király, I., Morin, A. J. S., Bőthe, B., Orosz, G., & Rigó, A. (2018). Investigating the multidimensionality of need fulfillment: A bifactor exploratory structural equation modeling representation. *Structural Equation Modeling*, 25, 267–286. <http://dx.doi.org/10.1080/10705511.2017.1374867>
- Tóth-Király, I., Orosz, G., Dombi, E., Jagodics, B., Farkas, D., & Amoura, C. (2017). Cross-cultural comparative examination of the Academic Motivation Scale using exploratory structural equation modeling. *Personality and Individual Differences*, 106, 130–135. <http://dx.doi.org/10.1016/j.paid.2016.10.048>
- Veneziani, C. A., Fuochi, G., & Voci, A. (2017). Self-compassion as a healthy attitude toward the self: Factorial and construct validity in an Italian sample. *Personality and Individual Differences*, 119, 60–68. <http://dx.doi.org/10.1016/j.paid.2017.06.028>

- Williams, M. J., Dalgleish, T., Karl, A., & Kuyken, W. (2014). Examining the factor structures of the five facet mindfulness questionnaire and the self-compassion scale. *Psychological Assessment, 26*, 407–418. <http://dx.doi.org/10.1037/a0035566>
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34*, 806–838. <http://dx.doi.org/10.1177/0011000006288127>
- Yarnell, L. M., & Neff, K. D. (2013). Self-compassion, interpersonal conflict resolutions, and well-being. *Self and Identity, 12*, 146–159. <http://dx.doi.org/10.1080/15298868.2011.649545>
- Young, K. S., van der Velden, A. M., Craske, M. G., Pallesen, K. J., Fjorback, L., Roepstorff, A., & Parsons, C. (2017). The impact of mindfulness-based interventions on brain activity: A systematic review of functional magnetic resonance imaging studies. *Neuroscience and Biobehavioral Reviews, 84*, 424–433.
- Zeng, X., Wei, J., Oei, T. P., & Liu, X. (2016). The Self-Compassion Scale is Not Validated in a Buddhist Sample. *Journal of Religion and Health, 55*, 1–14. <http://dx.doi.org/10.1007/s10943-016-0205-z>

Received June 12, 2017

Revision received May 3, 2018

Accepted May 18, 2018 ■