COMMENTARY

Commentary on Muris and Otgaar (2020): Let the Empirical Evidence Speak on the Self-Compassion Scale

Kristin D. Neff [1]

Abstract

This commentary addresses Muris and Otgaar's (2020) paper titled "The Process of Science: A Critical Evaluation of More than 15 Years of Research on Self-Compassion with the Self-Compassion Scale." The Self-Compassion Scale (SCS) is a multidimensional scale measuring self-compassion that includes subscales representing increased compassionate self-responding (CS)—self-kindness, common humanity, and mindfulness—and reduced uncompassionate self-responding (UCS)—self-judgment, isolation, and overidentification. Muris and Otgaar have proposed that a total SCS score should not be used because CS and UCS are separate and unrelated constructs. I propose that CS and UCS are distinct but related and that a total SCS score can validly be used. More than just asserting their viewpoint, however, Muris and Otgaar (2020) have made ad hominem attacks on any researchers who disagree with them as unscientific, irrational, or unethical. They claimed their position was irrefutable and that continued use of a total SCS score by researchers was evidence of bias. Although they acknowledged the need for empirical data confirming or disconfirming hypotheses generated by the different positions, they offered no confirmable hypotheses of their own and simply assumed the evidence supports their position. I lay out a series of ten hypotheses based on each position to test which is correct. The empirical evidence confirms the position that CS and UCS are related and that a total SCS score can be used, while disconfirming the position that CS and UCS are unrelated and that a total SCS score cannot be used.

Keywords Self-Compassion Scale · Self-compassion · Scientific bias · Psychometric validity

Muris and his Otgaar (2020) have published a paper in *Mindfulness* titled "The Process of Science: A Critical Evaluation of More than 15 Years of Research on Self-Compassion with the Self-Compassion Scale." This critique questions the validity of the field of self-compassion research due to problems posited concerning the Self-Compassion Scale (SCS; Neff, 2003b), which is used in the majority of studies on self-compassion. The SCS is a multidimensional scale measuring self-compassion that includes three subscales representing compassionate self-responding (CS)—self-kindness, common humanity, and mindfulness—and three subscales representing uncompassionate self-responding (UCS)—self-judgment, isolation, and overidentification. UCS items are reverse coded to indicate their absence. The SCS can be used to measure the six components separately, or else they can be combined into a total

score to measure the global mindset of self-compassion given that the components interact as a system. Muris and Otgaar (2020) argue that the SCS should not include subscales measuring UCS because they believe CS and UCS are separate and unrelated constructs. They assert that CS represents protection against psychopathology and UCS represents vulnerability to psychopathology and that they should not be examined as a whole.

Muris and his colleagues have written many critiques of the SCS (Muris, 2016; Muris, Otgaar, Meesters, et al., 2019a; Muris et al., 2016; Muris, Otgaar, & Pfattheicher, 2019b; Muris & Petrocchi, 2017; Muris et al., 2018). I have written many rebuttals to his critiques (Neff 2016a, 2016b, 2019) and have provided an abundance of empirical evidence that counters his claims (e.g., Neff, Long et al., 2018a; Neff et al., 2019). What is striking about this new offering by Muris and Otgaar (2020) is the patronizing and personal nature of their attacks. "It is apparent that there is irrefutable proof that the current conceptualization of self-compassion and the way this trait is currently assessed with the SCS are inappropriate." These two scholars appear bewildered by the fact that after all

✉ Kristin D. Neff
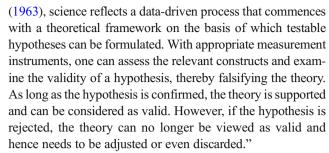   kneff@austin.utexas.edu

[1] Educational Psychology, The University of Texas at Austin, 1912 Speedway, Suite 504, Austin, TX 78712-1289, USA

of their criticisms of the SCS, I have still not submitted to their more rational viewpoint nor have I changed my theory or measurement of self-compassion. How is it possible that researchers such as myself are able to "maintain a certain perspective that is plain wrong and no longer in keeping with the main premises of their theory?" Why might this occur? "The short answer is that ... science is conducted by human beings who do not always operate in a logical, rational way but rather are driven by personal interests, cognitive biases, and social influences."

Muris and Otgaar (2020) not only questioned my rationality but that of the hundreds of other researchers who have used a total SCS score in their studies. They claimed "In the past years, multiple researchers have put effort in demonstrating that compassionate and uncompassionate self-responding components included in the SCS are dissimilar and do not represent a single protective trait (e.g., Brenner et al. 2017, 2018; Lopez et al. 2015; Muris et al. 2018, 2019a). That is, various types of methods were used to substantiate critiques regarding the scale (i.e., meta-analysis, face validity checks, empirical research), but so far, this has not led to a notable change in the way the SCS has been employed. Of course, we do not want to call into question the autonomy and self-governance of researchers, but the fact is that scientists, like all other people, are prone to various kinds of biases."

They went on to consider all the biases that were likely driving what they perceive as the clearly irrational and unscientific use of a total SCS scale score such as anchoring (i.e., relying too heavily on the first piece of information on a subject), confirmation bias (i.e., the tendency to search for, interpret, focus on, and remember information in a way that confirms one's preconceptions), and status quo bias (the tendency to want things to stay the same). These biases were thought to be driving the behavior of peer reviewers and journal editors as well, who were foolish enough to actually publish research that uses a total SCS score. Muris and Otgaar (2020) were not content with labeling the entire field of self-compassion researchers who do not agree with them as biased, however. They actually suggested that many are unethical: "Scientists are also human in the sense that they consciously behave in a way that is not in line with ethical regulations. Science can be a hypercompetitive activity requiring researchers to publish a good quantity of papers in high-impact journals and to acquire research grants, thereby creating a climate that may tempt them to cut corners, exaggerate findings, and overstate the importance of their research."

Muris and Otgaar appear to be so convinced that they are right and that everyone else is wrong, that the only possible explanation for the fact that others disagree with them must be a lack of rationality or ethics. They do not appear to even consider the possibility that our disagreement is based on empirical evidence. They instead assume that it stems from a disregard for the scientific process: "According to Popper (1963), science reflects a data-driven process that commences with a theoretical framework on the basis of which testable hypotheses can be formulated. With appropriate measurement instruments, one can assess the relevant constructs and examine the validity of a hypothesis, thereby falsifying the theory. As long as the hypothesis is confirmed, the theory is supported and can be considered as valid. However, if the hypothesis is rejected, the theory can no longer be viewed as valid and hence needs to be adjusted or even discarded."

Despite discussing the need for clear hypotheses that can confirm or disconfirm theory, Muris and Otgaar (2020) did not actually formulate a single testable hypothesis in their paper that was evaluated according to empirical evidence. Rather, they simply assumed the evidence supports their position and that the reason I have not changed my position is because I have ignored the relevant evidence.

I have taken the liberty, therefore, of creating clear testable hypotheses that would either confirm their position (that CS and UCS are distinct and unrelated and a total SCS scale score should not be used) or my own (all six subscales of the SCS, including those representing CS and UCS, are distinct but interrelated and can be used as a total score). I then present empirical evidence to evaluate these hypotheses. I believe that empirical evidence does matter and that the hundreds of researchers who have used a total SCS score are not deluded fools, but rather have the intelligence to judge the empirical evidence accurately. In my opinion, the evidence speaks for itself.

My Model of Self-compassion

Self-compassion is a healthy way of relating to the suffering that stems from feelings of inadequacy or general life challenges. I conceptualize self-compassion as a system comprised of six distinct but interrelated elements. Self-kindness entails being warm, supportive, and understanding towards oneself. Common humanity involves recognizing the shared human experience of imperfection, understanding that all humans fail and make mistakes. Mindfulness involves being aware of one's present moment experience of suffering with equanimity and balance. Self-judgment entails harshly criticizing oneself for one's failings. Isolation involves feeling abnormal and alone in the experience of suffering. Overidentification occurs when one is fused with one's suffering to the point that perspective is lost. The various components of self-compassion are conceptually distinct and tap into different ways that individuals emotionally respond to suffering (with more kindness and less judgment), cognitively understand suffering (as part of the human experience rather than as isolating), and pay attention to suffering (in a more mindful and less overidentified manner). Self-compassion involves simultaneously engaging in compassionate self-responding and disengaging in uncompassionate self-responding. I propose

that self-compassion can be measured using a total SCS score (representing self-compassion as a whole) or with six subscales (representing its six constituent components).

I suggest that when scholars want to understand the mechanisms of action of self-compassion in terms of how it impacts outcomes (e.g., how learning self-compassion impacts body image), that use of the six subscales is preferred. I recommend using the six subscales as opposed to global CS and UCS scores not only because two global factors are not psychometrically supported (see below) but because it provides more precise and nuanced information. A global CS score lumps together self-kindness, common humanity, and mindfulness, and a global UCS score merges self-judgment, isolation, and overidentification, obscuring key difference between the individual components. For example, Körner et al. (2015) examined the link between the six self-compassion subscales and depression in a large community sample using regression analyses and found that isolation predicted 18% of the variance in depressive symptomology, followed by overidentification and self-kindness which each predicted 2%, and mindfulness and self-judgment, which each predicted 1%. The use of the subscales helps to illuminate how the components differentially impact outcomes.

However, I suggest that when scholars want to understand the general relationship between a self-compassionate mindset and well-being, especially in terms of understanding how teaching people to be more self-compassionate might impact well-being, it is most parsimonious to use a total score representing self-compassion as a whole (Neff et al., 2019; Neff & Tóth-Király, 2020). This is especially because all six components cohere as a balanced system within individuals and change in tandem after therapeutic intervention (see below).

What is the Significance of Findings that UCS Predicts Psychopathology More Strongly than CS?

The pillar of Muris and colleagues' arguments against the use of a total SCS scale score is the fact that the UCS subscales tend to have a stronger relationship with psychopathology than the CS subscales (e.g., Brenner et al., 2018; Muris et al., 2019b; Muris & Petrocchi, 2017). Surely, this is solid empirical evidence disconfirming my model, is it not? No, it is not.

In my original theory paper (Neff, 2003b), I argued that it is the decrease in self-judgment, feelings of isolation, and over-identification with suffering entailed by a self-compassionate mindset that is likely responsible for the reduction of psychopathology associated with self-compassion. This is because reduced resistance to pain lessens suffering (Hayes et al., 1999). Findings that the UCS components are more strongly associated with pathological states such as depression, anxiety, and stress than the CS components are fully in line with my model.

While all six components of self-compassion significantly contribute to well-being, my research confirms that reductions in psychopathology appear to be more strongly driven by the lessened UCS that occurs in a self-compassionate mindset (Neff et al., 2018a). Whereas the CS subscales typically have small to medium correlations with psychopathology, the UCS subscales typically have medium to large correlations. This is why I recommend that when researchers are interested in understanding the mechanisms by which self-compassion impacts well-being, it is preferable to use subscale scores rather than a total score.

It is not clear why Muris and colleagues have assumed that because the UCS subscales predict psychopathology more strongly than the CS subscales, that these findings invalidate the use of a total SCS score. Their position does not accord with any known principles of scale validation in the field of psychometric studies. No psychometric scholar that I am aware of has argued that all of the subscales of a multidimensional measure must have the same strength of association with outcomes in order to justify the use of a total scale score, nor have Muris and colleagues cited any such arguments. In fact, it does not make sense to have a multidimensional scale in the first place if all subscales predict outcomes exactly the same way. If it were the case that the SCS was theorized to be a unidimensional scale, the fact that UCS items predict psychopathology more strongly than CS items would indeed be a problem. But, that is not the case. The SCS is multidimensional. The primary reason for having a multidimensional scale is because the subscales are distinct and presumably have differential associations with outcomes. Most commonly used multi-dimensional scales such as the Five Factor Mindfulness Questionnaire (Baer et al., 2006) or the Difficulties in Emotion Regulation Scale (Gratz & Roemer, 2004) contain subscales which display a differential pattern of association with outcomes. This actually helps to buttress the validity of the scale because it supports the inclusion of separate subscales, demonstrating its multi-dimensionality.

If Muris and Otgaar (2020) had simply urged scholars to use subscale scores rather than a SCS total score when investigating how self-compassion works to change outcomes, I would not have a problem with this position. This is not what they said, however. Rather, they asserted that it is incorrect and invalid to use a total SCS score under any circumstances, without offering any scientifically sound reason why not.

Are SCS Items a Face Valid Measure of Protection?

Muris and Otgaar believe that CS and UCS are distinct and unrelated processes, with the former indicating protection against psychopathology and the latter indicating vulnerability to psychopathology. It is quite right that CS items represent protection and UCS items represent vulnerability, but it is odd
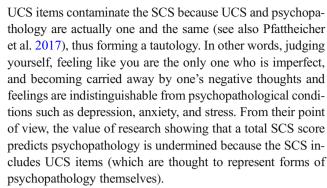
indeed to argue that protection and vulnerability are unrelated. Protection is defined as being shielded from harm. Vulnerability is defined as being exposed to harm. Protection, by definition, reduces vulnerability. Self-compassion protects against psychological harm through increased self-kindness, common humanity, and mindfulness and reduced self-judgment, isolation, and overidentification. CS and reduced UCS go hand in hand.

Muris and Otgaar suggest that because self-compassion is said to offer protection against psychopathology, the fact that UCS items represent vulnerability undermines the face validity of the scale. However, they have failed to take into account the fact that UCS items are reverse coded to indicate their absence. Items are written in a straightforward manner and subsequently reverse coded in order to avoid double negatives. For instance, it is easier to respond on a scale from "almost never" to "almost always" to the item "I'm disapproving and judgmental about my own flaws and inadequacies" (which is reverse coded) than it is to respond to the item "I am not disapproving and judgmental about my own flaws and inadequacies."

Muris et al. (2018) conducted a study to assess the face validity of the SCS by giving SCS items to psychologists and asking them to judge whether each item represented healthy protection or unhealthy vulnerability. Unsurprisingly, they found that psychologists judged the CS items to represent healthy protection and UCS items to represent unhealthy vulnerability. But, they forgot to do the very obvious—make it clear to respondents that the items representing USC are reverse coded to indicate their absence. A scientifically sound comparison of our two models with regard to the face validity of the SCS would involve asking therapists to classify whether a response of "almost never" to UCS items such as "When I fail at something important to me I become consumed by feelings of inadequacy" is an indication of healthy protection or unhealthy vulnerability. Responses indicating healthy protection would support my point of view, responses indicating unhealthy vulnerability would support Muris' point of view. It is almost certain that therapists would have classified the absence of UCS as indicating healthy protection. How could it possibly be otherwise? Would any licensed therapist actually classify an individual who is not consumed by feelings of inadequacy as demonstrating unhealthy vulnerability rather than healthy protection? This is the quality of the scientific evidence Muris and Otgaar (2020) have asserted is so irrefutable that anyone who is not swayed by it must be unscientific, irrational, and suffering from unconscious biases.

Is UCS the Same as Psychopathology?

Muris and Otgaar (2020) asserted that the SCS is "contaminated by psychopathological characteristics." They believe UCS items contaminate the SCS because UCS and psychopathology are actually one and the same (see also Pfattheicher et al. 2017), thus forming a tautology. In other words, judging yourself, feeling like you are the only one who is imperfect, and becoming carried away by one's negative thoughts and feelings are indistinguishable from psychopathological conditions such as depression, anxiety, and stress. From their point of view, the value of research showing that a total SCS score predicts psychopathology is undermined because the SCS includes UCS items (which are thought to represent forms of psychopathology themselves).

While it is true that UCS is associated with psychopathology, they are clearly not the same thing. Psychopathological conditions like depression, anxiety, and stress are generalized mood disorders or types of mental illness that impair functioning. For example, items from the Depression, Anxiety and Stress Scale (Lovibond & Lovibond, 1995) measure the degree to which individuals experience depression (e.g., "I found it difficult to work up the initiative to do things"), anxiety (e.g., "I experienced breathing difficulty... excessively rapid breathing, breathlessness in the absence of physical exertion"), and stress (e.g., "I found it hard to wind down."). In contrast, the three UCS subscales are ways of relating to personal experiences of suffering: harshly judging one's inadequacies, believing one is unique from others, and ruminating on negative emotions. Psychopathology may sometimes be caused by UCS, but it may also be caused by hormonal imbalances, current life stressors, or other factors. Similarly, UCS does not always lead to mental illness. UCS and psychopathology are not synonymous. Importantly, you cannot directly teach people how to be less depressed, anxious, or stressed. However, you *can* change the circumstances that lead to these mood disorders. You can put people on medication, you can reduce their life stressors, or you can help them learn to be more self-compassionate (i.e., simultaneously increase CS and decrease UCS).

Muris and Otgaar (2020) offered no empirical evidence to support the hypothesis that UCS and psychopathology are one and the same beyond citing studies showing that they are correlated (e.g., Brenner et al., 2018; Muris & Petrocchi, 2017). Correlation does not indicate sameness. A standard way to test whether constructs are the same or different is to establish incremental validity—in other words, to test whether one construct predicts a specified outcome over and above another predictor. In order for the claim that the two constructs are undifferentiated or tautological to be supported, the empirical evidence would need to demonstrate an absence of incremental validity. In response to Pfattheicher et al. (2017), who made a similar argument that UCS items are indistinguishable from neuroticism, Neff et al. (2018b) conducted three studies to test the incremental validity of UCS items with neuroticism using the NEO-PRI (Costa & McCrea, 1992). This scale measures six facets of neuroticism: anxiety, depression,

vulnerability, hostility, self-consciousness, and impulsiveness. The UCS items predicted variance over and above neuroticism as a whole and also the depression and anxiety facets in particular. The first study established incremental validity with regard to life satisfaction, the second with difficulties in emotion regulation, and the third with wisdom, happiness, optimism, self-esteem, psychological well-being, and negative affect.

The available empirical evidence therefore does not support the hypothesis that UCS and psychopathology are the same, but rather supports the hypotheses that they are different. Muris and Otgaar (2020) ignored this empirical data, however, and simply asserted that because UCS items are correlated with psychopathology they should be dropped from the SCS. Using this line of reasoning, any items in any psychological measure that correlate with psychopathology should be thrown out because they are tautological. Such a position would not, of course, be tenable.

How Does Self-compassion Change Through Training?

If Muris and Otgaar are right in their proposition that self-compassion involves CS but not reduced UCS and that the two are unrelated, self-compassion training should hypothetically increase CS more than it reduces UCS. If CS and UCS are interrelated and form a system, however, they should increase to roughly the same degree. Ferrari et al. (2019) conducted a recent meta-analysis of 27 randomized controlled trials of self-compassion interventions and found that all six subscales of the SCS change significantly as a result of training and that the strongest effect size was actually found for reductions in overidentification. In a study of the Mindful Self-Compassion program (MSC; Neff & Germer, 2013), changes in the subscales were strikingly similar: self-kindness increased 36% and self-judgment decreased 32%; common humanity increased 34% percent and isolation decreased 35%; mindfulness increased 21% and overidentification decreased 33% (Neff, 2016a). These data directly disconfirm Muris and Otgaar's position.

This is important because one of their arguments is that the use of a total SCS scale score will lead clinicians astray when deciding whether to use self-compassion as a treatment modality. If UCS is a stronger predictor of psychopathology than CS, is it worth it to teach clients self-compassion? The answer is yes because teaching them self-compassion will reduce UCS. The whole point of compassion-based therapies such as Compassion-Focused Therapy (CFT; Gilbert 2009, 2014) is that learning to be more self-compassionate reduces the tendency to judge oneself, to feel isolated and disconnected from others, or be become absorbed in negative emotions such

as shame. Unlike Cognitive Behavioral Therapy, which often targets negative self-talk directly (Hofmann et al., 2012), therapies such as CFT focus on helping clients build skills of self-soothing as a means to reduce self-criticism.

Muris and Otgaar (2020) are concerned that the use of a total SCS scale score will make it more difficult to ascertain the particular changes that occur in therapy that lead to well-being. I agree. This is why I recommend that researchers use all six subscales when wanting to examine the mechanisms of self-compassion. But if I am a clinician who simply wants to know whether or not teaching my clients self-compassion will help them be less depressed, the fact that self-compassion as a whole is linked to less depression is more relevant than knowing that the CS or UCS subscales differentially impact outcomes. As a clinician, there is no need to target CS and UCS separately because they change in tandem. Moreover, it is more effective to teach clients what to do (be kinder, remember common humanity, and be mindful) than what not to do (be less judgmental, feel less isolated, stop overidentifying with suffering). The result of being more compassionate is being less uncompassionate. This is the foundation on which therapies such as CFT and training programs such as MSC are based.

Does the SCS Inflate the Link Between Self-compassion and Psychopathology?

Muris and Otgaar (2020) have claimed that the SCS "inflates" the link between self-compassion and psychopathology, but have provided absolutely no empirical evidence to support this claim other than showing that the UCS subscales are more strongly predictive of psychopathology than the CS subscales (e.g., Brenner et al., 2018; Muris & Petrocchi, 2017). To empirically evaluate whether the SCS inflates the link with psychopathology, it is necessary to examine whether research using a total SCS score demonstrates a stronger negative relationship between self-compassion and psychopathology than occurs in real-life therapeutic contexts. Evidence that much stronger relationships are found using a total SCS score than is found in real-life contexts would support Muris and Otgaar's position, while findings that the SCS accurately represents this relationship would support my own. A total SCS score typically has medium to strong negative correlations with stress, depression, and anxiety (Neff et al., 2018a) and the meta-analysis conducted by Ferrari et al. (2019) also found moderate to strong effects for reduced stress, depression, and anxiety as a result of self-compassion interventions. Once again, the empirical evidence disconfirms Muris and Otgaar's position and confirms my own.

What Does the Psychometric Research Say?

Muris and Otgaar (2020) dismissed the research conducted on the SCS to support its validity (e.g., Cleare et al., 2018; Neff et al., 2019; Neff et al., 2018b; Neff et al., 2017) as a "scientific smoke curtain." They argued that "the exact factor structure of the SCS is far from clear and tends to differ across studies. It is important to note though that all studies have in common that a simple one-factor model does never provide the best fit for the SCS." This statement is very telling and reveals a woeful lack of understanding of the theoretical and psychometric issues at hand. I have never proposed that self-compassion is unidimensional and have argued from the beginning (Neff, 2003a, 2003b) that it is a multidimensional construct that needs to be measured with a multidimensional scale. Moreover, many of the psychometric studies challenging the factor structure of the SCS have been incomplete and deeply flawed. These studies typically use methods of analysis that cannot adequately model multidimensional constructs that form a system. For instance, Lopez et al. (2015) used exploratory factor analyses to find a two-factor solution to the SCS, even though EFA cannot explicitly be used in a confirmatory manner, nor can it model a complex system. Brenner et al. (2017) found that two general factors had a better fit for the SCS than one general factor using CFA bifactor analyses, but examined two uncorrelated general factors, which is at odds with self-compassion theory. Halamová et al. (2020) found two factors for the SCS using item response theory (IRT), but IRT is designed for unidimensional rather than multidimensional constructs, rendering findings meaningless. Instead, more sophisticated and complex psychometric methods are needed to model systems.

The bifactor ESEM framework (Morin et al., 2016) is an increasingly popular method used by many scholars to validate multidimensional measures (Boateng et al., 2018). The advantage of this framework is that it can model two sources of construct-relevant multidimensionality—the simultaneous operating of general and specific factors, and the loading of items on multiple specific factors—for a more precise psychometric examination of multidimensional measures.

In conjunction with 17 international colleagues (Neff et al., 2019), we relied on this framework in the largest and most comprehensive study of the factor structure of the SCS to date. The study was designed to directly test whether the factor structure of the SCS is best represented by a single or two global factors representing CS and UCS. If Muris and Otgaar's position was correct, CS and UCS should form two factors but not one general factor, and if my position was correct, items should form one general factor (with six specific factors). We examined this question in 20 diverse international samples ($N = 11,685$). Five different models using both CFA and ESEM were systematically tested and compared: a one-factor, two-factor correlated (representing CS and UCS), six-factor correlated (representing the six elements), a bifactor model (representing one general self-compassion factor and six specific factors), and a correlated two-bifactor model (representing two general CS or UCS factors each with three specific factors).

The data disconfirmed Muris and Otgaar's position and confirmed my own. In every single sample examined, the bifactor ESEM model had excellent fit and the global self-compassion factor was defined by strong factor loadings. Moreover, results indicated that 95% percent of the reliable variance in item responding was explained by a general factor of self-compassion. In contrast, a correlated two-factor model and the correlated two-bifactor model using CFA had poorer fit in every sample examined compared with the bifactor ESEM model. The correlated two-bifactor model using ESEM had a good fit, but factor loadings indicated that this solution was poor: Positive and negative items did not load strongly on the positive and negative factors, meaning that CS and UCS could not be empirically differentiated. Results supported the use of one rather than two global factors.

A recent study by Tóth-Király and Neff (2020) demonstrated that the factor structure of the SCS is invariant across culture, gender, age, and population type (e.g., student, community, or clinical) in 18 international samples. Moreover, the same factor structure was supported in an adaptation of the SCS for youths (Neff, Bluth et al., 2020a), an adaptation focused on compassion for others (Pommier et al., 2020), and a state version of the SCS (Neff, Tóth-Király et al., 2020b). This suggests that the bifactor ESEM framework provides a stable way to model the factor structure of compassion.

It is important to note that we have typically examined the models proposed by critics in our psychometric papers, including a two-factor model and a two-bifactor correlated model using both CFA and ESEM so that we could empirically compare and contrast proposed models (Neff et al., 2018b; Neff et al., 2019; Neff et al., 2020a; Neff et al., 2020b). However, there is not a single study conducted by those who argue that the SCS forms two factors rather than one factor that has included our proposed bifactor ESEM model. In every one of the samples we have analyzed that have compared one versus two global factors (over 26 and counting), it was demonstrated that a single bifactor ESEM model performs the best. Although it will be important that other scholars replicate our findings with the SCS, the current psychometric evidence is quite clear in supporting the use of a global SCS score.

How is Self-compassion Configured Within Individuals?

The fact that 95% of the variance in SCS item responding can be explained by a total score provides extremely strong

evidence for use of a total SCS score and suggests that items representing CS and UCS cohere as a system. This evidence is not only apparent when examining pools of participants but also when examining how the various elements of self-compassion are configured within individuals. If Muris and Otgaar (2020) are correct in their assertion that CS and reduced UCS are distinct and unrelated, then hypothetically there should be many individuals who are high in both or low in both. If my position is correct that CS and UCS subscales are distinct but form a balanced system, most individuals should either be high in CS and low in UCS, low in CS and high in UCS, or moderate in both, and very few individuals should be high or low in both. Phillips (2019) used latent profile analyses in a sample of community adults and a second sample of undergraduates to examine profiles or patterns of scores on the various SCS subscales. She found only three patterns—high in the three CS subscales and low in the three UCS subscales, low in CS subscales and high in UCS subscales, or moderate in both. She did not find any individuals who were high or low in both. Ulrich-French and Cox (2020) performed similar analyses across three samples of undergraduates and found that the vast majority of individuals displayed one of these three patterns, although they also identified a tiny fraction (between 5 and 8% across studies) who were high or low in both UCS and CS. However, the researchers did not include any attention checks. Undergraduates often hurry to complete surveys by responding to all items the same way (Oppenheimer et al., 2009), so findings that a few individuals displayed an aberrant pattern may have been due to a method effect. Regardless, these individuals are uncommon exceptions that prove the rule. The empirical evidence suggests that CS and UCS form a balanced system within individuals, disconfirming Muris and Otgaar's position and confirming my own.

How Does the State of Self-compassion Operate?

One of the best ways to determine if self-compassion includes reduced UCS, and to understand whether or not self-compassion operates as a holistic system or if CS and UCS operate independently, is by examining self-compassion as a state. The SCS measures self-compassion as a trait—the habitual tendency to be in a state of self-compassion when confronted with various types of negative experiences such as feelings of personal inadequacy or general life difficulties. There is bound to be error in the measurement of self-compassion when examined as a trait due to variance in the types of situations individuals are considering across items. We have recently created the State Self-Compassion Scale (SSCS) to examine how self-compassion operates in real time (Neff et al., 2020b). The SSCS asks participants to think of particular difficulty or instance of suffering and then respond to all items with reference to how they are relating to that particular difficulty. If Muris and Otgaar's position is correct and CS and UCS are distinct and unrelated constructs, then hypothetically state CS and UCS items should form two separate factors. If my position is correct CS and UCS components are distinct and interrelated, items should form a single general factor (with the six elements as specific factors). The evidence disconfirms the position of Muris and Otgaar and confirms my own. Similar to the trait SCS, the factor structure of the SSCS is best represented by a bifactor ESEM model with one general and six specific factors, rather than two global factors. It is also important to note that the correlation between CS and reduced UCS was extremely high across two studies for the two-bifactor CFA model ($r = .887$ in study 1 and $r = .833$ in study 2), which is higher than has typically been found in research with the trait SCS (e.g., Coroiu et al., 2018; Costa et al., 2016; López et al., 2015; Pfattheicher et al., 2017). This finding is likely due to the fact that all items were aimed at the same instance of suffering and therefore assessed the experience of self-compassion itself as opposed to reflecting variance stemming from responding to different types of situations. Clearly, CS and UCS do not operate independently.

Does Inducing a Self-compassionate Mindstate Change CS More than UCS?

As part of the validation process for the SSCS (Neff et al., 2020b), we also gave participants a task designed to induce a self-compassionate mindstate by writing about a difficult situation with mindfulness, common humanity, and kindness (adapted from Leary et al., 2007). We then examined how state self-compassion in relation to that situation changed as a result. If Muris and Otgaar's position is correct and the construct of self-compassion does not include reduced UCS, then hypothetically state CS should change more than state UCS after a self-compassion mindstate induction. If my position is correct, CS and UCS should change in tandem. We found the manipulation created change in the six components of state self-compassion to a remarkably similar degree: self-kindness increased 10.4% and self-judgment decreased 10.4%; common humanity increased 11.6% percent and isolation decreased 11.4%; mindfulness increased 9.2% and overidentification decreased 8.4%. These findings once again disconfirm Muris and Otgaar's position and confirm my own. The fact that this system change could be observed experimentally is especially important given that the change occurred at the moment, focused on the same experience of suffering.

Table 1 Empirical evidence related to hypotheses confirming or disconfirming Muris and Otgaar's position

Hypotheses	Confirm	Disconfirm
1. UCS more strongly predicts psychopathology than CS[a]	N/A	N/A
2. UCS items are not face valid indicators of the protective function of self-compassion[b]	--	--
3. UCS should not demonstrate incremental validity with measures of psychopathology		X
4. Self-compassion training should change scores on CS items more than UCS items		X
5. The SCS should demonstrate a stronger link with psychopathology than is observed in real life contexts		X
6. SCS items should not form a global factor in psychometric analyses		X
7. Some individuals should be high in both CS and UCS or low in both CS and UCS		X
8. Items measuring state self-compassion should not form a global factor in psychometric analyses		X
9. Experimental manipulation of self-compassion should change state CS items more than state UCS items		X
10. CS and UCS should be differentially associated with sympathetic and parasympathetic nervous system activity and brain function		X

[a] This hypothesis is not applicable to testing the position that CS and UCS are unrelated or whether a total SCS scale score should be used

[b] The available empirical evidence is not sufficient to test the hypothesis because it does not account for reverse coding

What Does the Physiological Evidence Show?

One of the explanations Muris and Otgaar (2020) provided for the fact that researchers continue to use the SCS as a total score despite what they believe is obvious and "irrefutable" evidence to the contrary, is that there is not a viable alternative model available. However, they note that Paul Gilbert's Social Mentality Theory (Gilbert, 2000, 2005) proposes that CS involves the parasympathetic nervous system associated with care and attachment and that UCS involves the sympathetic nervous system associated with threat-defense. "Thus, there is a solid alternative theory...proposing that compassionate and uncompassionate self-responding indeed reflect different processes that are moderated by different brain systems...which—as noted earlier—implies that it is not appropriate to combine them in a single score of self-compassion."

If Muris and Otgaar's point of view were to be supported with physiological evidence using Gilbert's model, hypothetically the CS subscales should be more strongly associated with parasympathetic nervous system activation and the UCS subscales should be more strongly associated with sympathetic activation. They should also be differentially associated with brain function. In fact, the CS and UCS subscales are not differentially associated with physiological markers of sympathetic activity such as alpha-amylase or interleukin-6 (Neff et al., 2018a) or with markers of parasympathetic activity such as heart rate variability (Svendsen et al., 2016). Similarly, Parrish et al. (2018) found that CS and UCS subscales did not differentially predict ventromedial prefrontal cortex and amygdala connectivity during a social feedback task. Thus, the physiological evidence disconfirms Muris and Otgaar's position and confirms my own.

Gilbert and his colleagues have conducted related research examining brain activation during the process of reassuring or criticizing oneself (similar to CS and UCS). An fMRI study by Longe et al. (2009) asked participants to imagine responding with self-reassurance or self-criticism to a negative social situation and examined their brain activity. Findings indicated that self-criticism was associated with activity in the lateral prefrontal cortex and dorsal anterior cingulate regions, while self-reassurance activated the left temporal pole and insula. Gilbert and colleagues have used these findings to argue that CS and UCS subscales should not be combined in a total SCS score because the brain activity associated with these processes is distinct (Gilbert et al., 2011). However, a more recent study conducted by Gilbert and his colleagues (Kim et al., 2020) also used fMRI imagery to examine reactions to negative emotional stimuli. This study found that self-criticism increased activity in the anterior insula, anterior cingulate, and the amygdala and that self-reassurance suppressed activity in these very same regions. The findings of the latter study provide strong support for the view that CS and UCS are interrelated rather than independent. The findings of Longe et al., 2009 do not disconfirm my model, however, because I do not argue that CS and UCS components are the same, just that they are interrelated.

Ultimately, the debate about the conceptualization and measurement of self-compassion will be decided based on the empirical evidence. Table 1 presents ten hypotheses for evaluating the position of Muris and Otgaar (CS and UCS

are distinct and unrelated and a total SCS score cannot be used) and my own (CS and UCS are distinct but interrelated and a total SCS score can be used). Hypothesis 1 reflects the primary reason Muris and Otgaar believe a total SCS score should not be used: UCS items are more strongly related to psychopathology than CS items. This hypothesis is irrelevant to evaluating the use of a total SCS score, however, because the SCS is a multidimensional measure and it is common for distinct subscales to differentially predict outcomes in such measures. These findings are also in full accord with my model of how self-compassion works (Neff, 2003b). Hypothesis 2 concerns the face validity of UCS items but cannot be properly evaluated given that relevant studies have not accounted for reverse coding. (It is almost certain that reverse-coded UCS items would be categorized as representing healthy protection, disconfirming Muris and Otgaar's position.) Empirical evidence for the other eight hypotheses clearly supports my position and disconfirms that advocated by Muris and Otgaar (2020).

The available empirical evidence indicates that CS and reduced UCS operate in tandem and that use of a total SCS scale score is valid. I hope that Muris and Otgaar can adhere to the Popperian principles they espouse: "As long as the hypothesis is confirmed, the theory is supported and can be considered as valid. However, if the hypothesis is rejected, the theory can no longer be viewed as valid and hence needs to be adjusted or even discarded." Even if they do not, I would hope at the very least that they will refrain in the future from ad hominem attacks and calling those who disagree with their point of view unscientific, biased, and unethical. That, indeed, is truly unscientific.

Compliance with Ethical Standards

Conflict of Interest   The author declares that she has no conflict of interest

Ethical Approval   The manuscript does not involve human or animal participants.

References

Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J., & Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment, 13*(1), 27–45.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*(149). https://doi.org/10.3389/fpubh.2018.00149.

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

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

Cleare, S., Gumley, A., Cleare, C. J., & O'Conner, J. C. (2018). An investigation of the factor structure of the Self-Compassion Scale. *Mindfulness, 9*(2), 618–628.

Coroiu, A., Kwakkenbos, L., Moran, C., Thombs, B., Albani, C., Bourkas, S., Zenger, M., Brahler, E., & Korner, A. (2018). Structural validation of the Self-Compassion Scale with a German general population sample. *PLoS One, 13*(2), e0190771.

Costa Jr., P. T., & McCrae, R. R. (1992). *NEO-PI-R Professional Manual.* Odessa, FL: Psychological Assessment Resources.

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

Ferrari, M., Hunt, C., Harrysunker, A., Abbott, M. J., Beath, A. P., & Einstein, D. A. (2019). Self-compassion interventions and psychosocial outcomes: A meta-analysis of RCTs. *Mindfulness, 10*, 1455–1473.

Gilbert, P. (2000). Social mentalities: internal 'social' conflict and the role of inner warmth and compassion in cognitive therapy. In P. Gilbert & K. G. Bailey (Eds.), *Genes on the couch: Exploration in evolutionary psychotherapy* (pp. 118–150). Hove: Psychology Press.

Gilbert, P. (2005). *Compassion: Conceptualizations, Research, and Use in Psychotherapy.* London: Routledge.

Gilbert, P. (2009). Introducing compassion-focused therapy. *Advances in Psychiatric Treatment, 15*(3), 199–208.

Gilbert, P. (2014). The origins and nature of compassion focused therapy. *British Journal of Clinical Psychology, 53*(1), 6–41.

Gilbert, P., McEwan, K., Matos, M., & Rivis, A. (2011). Fears of compassion: Development of three self-report measures. *Psychology and Psychotherapy, 84*, 239–255.

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment, 26*(1), 41–54.

Halamová, J., Kanovský, M., Petrocchi, N., Moreira, H., López Angarita, A. et al. (2020). The factor structure of the Self-Compassion Scale in eleven distinct populations across the world. *Measurement and Evaluation in Counseling and Development.* Advance online publication. DOI: https://doi.org/10.1080/07481756.2020.1735203

Hayes, S. C., Strosahl, K., & Wilson, K. G. (1999). *Acceptance and Commitment Therapy: Understanding and treating human suffering.* New York: Guilford.

Hofmann, S. G., Asnaani, A., Vonk, I. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research, 36*(5), 427–440.

Kim, J. J., Parker, S. L., Doty, J. R., Cunnington, R., Gilbert, P., & Kirby, J. N. (2020). Neurophysiological and behavioural markers of compassion. *Scientific Reports, 10*:6789. https://doi.org/10.1038/s41598-020-63846-3 .

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

Leary, M. R., Tate, E. B., Adams, C. E., Batts Allen, A., & Hancock, J. (2007). Self-compassion and reactions to unpleasant self-relevant events: The implications of treating oneself kindly. *Journal of Personality and Social Psychology, 92*(5), 887–904.

Longe, O., Maratos, F. A., Gilbert, P., Evans, G., Volker, F., Rockliff, H., et al. (2009). Having a word with yourself: Neural correlates of self-criticism and self-reassurance. *Neuroimage, 49*, 1849–1856.

López, A., Sanderman, R., Smink, A., Zhang, Y., van Sonderen, E., Ranchor, A., & Schroevers, M. J. (2015). A reconsideration of the Self-Compassion Scale's total score: Self-compassion versus self-criticism. *PLoS One, 10*(7). https://doi.org/10.1371/journal.pone.0132940.

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy, 33*, 335–343.

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

Muris, P. (2016). A protective factor against mental health problems in youths? A critical note on the assessment of self-compassion. *Journal of Child and Family Studies, 25*, 1461–1465. http://self-compassion.org/wp-content/uploads/2015/12/Mosewich_intervention.pdf.

Muris, P., & Otgaar, H. (2020). The process of science: A critical evaluation of more than 15 years of research on self-compassion with the Self-Compassion Scale. *Mindfulness, 11*(6), 1469-1482. https://doi.org/10.1007/s12671-020-01363-0 .

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

Muris, P., Otgaar, H., & Petrocchi, N. (2016). Protection as the mirror image of psychopathology: Further critical notes on the self-compassion scale. *Mindfulness, 7*(3), 787–790.

Muris, P., van den Broek, M., Otgaar, H., Oudenhoven, I., & Lennartz, J. (2018). Good and bad sides of self-compassion: A face validity check of the self-compassion scale and an investigation of its relations to coping and emotional symptoms in non-clinical adolescents. *Journal of Child and Family Studies, 27*(8), 2411–2421.

Muris, P., Otgaar, H., Meesters, C., Heutz, A., & Van den Hombergh, M. (2019a). Self-compassion and adolescents' positive and negative reactions to daily life problems. *Journal of Child and Family Studies, 28*, 1433–1444.

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

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

Neff, K. D. (2003b). Self-compassion: An alternative conceptualization of a healthy attitude toward oneself. *Self and Identity, 2*, 85–102.

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

Neff, K. D. (2016b). Does self-compassion entail reduced self-judgment, isolation, and over-identification? A response to Muris, Otgaar, and Petrocchi (2016). *Mindfulness, 7*(3), 791–797.

Neff, K. D. (2019). Setting the record straight about the Self-Compassion Scale. *Mindfulness, 10*, 200–202.

Neff, K., & Germer, C. (2013). A pilot study and randomized controlled trial of the mindful self-compassion program. *Journal of Clinical Psychology, 69*(1), 28–44.

Neff, K. D., & Tóth-Király, I. (2020). Self-Compassion Scale (SCS). In O. N. Medvedev, C. U. Krägeloh, R. J. Siegert, & N. N. Singh (Eds.), *Handbook of assessment in mindfulness*. New York: Springer. In press.

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.

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

Neff, K. D., Tóth-Király, I., & Colosimo, K. (2018b). 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, 32*(4), 371–392.

Neff, K. D., Tóth-Király, I., Yarnell, L., Arimitsu, K., Castilho, P., Ghorbani, N., Guo, H. X., Hirsch, J., Hupfeld, J., Hutz, C., Kotsou, I., Lee, W. K., Montero-Marin, J., Sirois, F., de Souza, L., Svendsen, J., Wilkinson, R., & Mantios, M. (2019). Examining the factor structure of the Self-Compassion Scale using exploratory SEM bifactor analysis in 20 diverse samples: Support for use of a total score and six subscale scores. *Psychological Assessment, 31*(1), 27–45.

Neff, K. D., Bluth, K., Tóth-Király, I., Davidson, O., Knox, M. C., Williamson, Z., & Costigan, A. (2020a). Development and validation of the Self-Compassion Scale for Youth. *Journal of Personality Assessment* Advance online publication. https://doi.org/10.1080/00223891.2020.1729774.

Neff, K. D., Tóth-Király, I., Knox, M., Kuchar, A. & Davidson, O. (2020b). *Examining self-compassion experimentally: The development and validation of a state self-compassion scale and mindstate induction.* Manuscript submitted for publication.

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45*(4), 867–872.

Parrish, M. H., Inagaki, T. K., Muscatell, K. A., Haltom, K. E., Leary, M. R., & Eisenberger, N. I. (2018). Self-compassion and responses to negative social feedback: The role of fronto-amygdala circuit connectivity. *Self and Identity, 17*(6), 723–738.

Pfattheicher, G., Hartung, W., & Schindler, S. (2017). Old wine in new bottles? The case of self-compassion and neuroticism. *European Journal of Personality, 31*(2), 160–169.

Phillips, W. J. (2019). Self-compassion mindsets: The components of the Self-Compassion Scale operate as a balanced system within individuals. *Current Psychology* Advance online publication. https://doi.org/10.1007/s12144-019-00452-1.

Pommier, E., Neff, K. D., & Tóth-Király, I. (2020). The development and validation of the Compassion Scale. *Assessment, 127*(1), 21–39.

Popper, K. R. (1963). *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Routledge.

Svendsen, J. L., Osnes, B., Binder, P. E., Dundas, I., Visted, E., Nordby, H., et al. (2016). Trait self-compassion reflects emotional flexibility through an association with high vagally mediated heart rate variability. *Mindfulness, 7*(5), 1103–1113.

Tóth-Király, I., & Neff, K. D. (2020). Is self-compassion universal? Support for the measurement invariance of the Self-Compassion Scale across populations. *Assessment.* In press.

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