



The Process of Science: A Critical Evaluation of more than 15 Years of Research on Self-Compassion with the Self-Compassion Scale

Peter Muris^{1,2} · Henry Otgaar^{1,3}

© The Author(s) 2020

Abstract

Objectives Falsification is regarded as a cornerstone of science. Anomalous data—even when highly credible—do not always lead to an adjustment of theory. In this paper, we discuss reasons for why current theories are not modified despite a corpus of work falsifying (parts of) the theory, using the case of self-compassion as an example. During more than 15 years, this psychological construct has been heralded as a protective factor against stress and emotional adversity.

Methods A search in the Web of Science database using [SELF-COMPASSION in title] as the search term was conducted and found empirical papers were critically evaluated.

Results Good evidence abounds indicating that the theoretical definition of self-compassion is limited and that as a consequence its proper assessment with the Self-Compassion Scale (SCS) is obscured as the measure is contaminated by psychopathological characteristics. Surprisingly, few researchers take these critical notes into account when conducting and reporting new studies on this topic. We argue that this might be due to all kinds of (conscious and unconscious) tendencies and cognitive biases of researchers and clinicians.

Conclusion Research lines are not always guided by solid data but strongly determined by social factors. We close by providing a recommendation for researchers regarding the assessment of self-compassion including the continued use of the SCS.

Keywords Self-compassion · Process of science · Anomalous data · Cognitive biases

Science refers to a process that builds and structures knowledge in the form of testable explanations and predictions. In psychology, the comprehension of (human) behavior and mental processes is the main spearpoint of scientific inquiry, and in the case of clinical psychology, the main theme of research is psychopathology, its phenomenology, associated etiological mechanisms, and their amplification or abridgment through interventions. Science is highly dependent on the definition and operationalization of the constructs under investigation, and this is true for psychology and clinical psychology in particular, which are dominated by concepts such as emotion and cognition that have been conceptualized in a variety of ways (e.g., Barrett 2017; Bayne et al. 2019). The multimodal

definition of constructs in this branch of science may give rise to debate and controversy, which oftentimes—after substantial empirical inquiry—results in a coexistence of diverse perspectives (e.g., Toomela 2019). However, occasionally, it happens that even in spite of convincing evidence, researchers maintain a certain perspective that is plain wrong and no longer in keeping with the main premises of their theory.

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. By this view, science reflects a logical, rational process that is fully driven by acquired empirical data. Oftentimes, science indeed operates according to Popperian principles. Within the field of clinical psychology, examples can be found of theories that after appropriate scientific inquiry had to be refuted or corrected. A case in point has been the

✉ Peter Muris
peter.muris@maastrichtuniversity.nl

¹ Department of Clinical Psychological Science, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands

² Stellenbosch University, Stellenbosch, South Africa

³ Catholic University of Leuven, Leuven, Belgium

hyperventilation theory of panic, which assumed that the typical physiological arousal symptoms (e.g., palpitations, breathlessness, chest pain, dizziness) as seen in patients with panic disorder are produced by overbreathing producing a metabolic state of respiratory alkalosis (e.g., Ley 1985). Experimental studies, however, showed that not the hyperventilation-induced physical symptoms per se but the catastrophic *interpretations* of such symptoms (e.g., palpitations interpreted as a sign of an impending heart attack) are the vehicle behind the development of panic attacks (Hornsveld 1996; Salkovskis and Clark 1990). This work on the interpretation of such symptoms has made researchers abandon the hyperventilation theory of panic.

Meanwhile, it should be noted that science does not always progress in a logical and rational way, and oftentimes, new evidence that does not accord well with the original data is disputed and rejected, leading to a consolidation of a perhaps objectively non-valid theory. This process of hampering scientific progress is well described by Kuhn (1962) in *The Structure of Scientific Revolutions*. Kuhn noted that empirical findings that are in disagreement with a current theory (so-called anomalies) do not immediately lead to a change in theory but rather could result in a period of “crisis” during which new methods and approaches of inquiry are permitted. These new scientific endeavors could then result in a “paradigm shift”, a significant adjustment or even complete replacement of the original theory. Such shifts do not take place readily and automatically; they are obstructed by social factors such as scientists’ interest in publishing (preferably high-impact) papers, various types of cognitive biases such as confirmation bias which can lead researchers to continue to defend and refuse to let go of a theory in spite of non-fitting new evidence. If this unfolds, a research program might enter a degenerative phase which could ideally lead to an abandonment of a certain theory (Lakatos and Musgrave 1974).

Scholars have repeatedly noted that the choice of observation and choice of attention are key to science, which means that researchers can be quite selective in what comes into their scientific lens and investigation (e.g., McComas 1996; Schwartz et al. 2004). Sometimes, however, there is undeniable proof that a theory is wrong or that the way a theory is currently investigated is seriously flawed. In our view, the latter is truly the case with the research program of self-compassion, a concept that was introduced more than 15 years ago as part of the positive psychology wave in clinical psychology. This research program has fueled the development of self-compassion related interventions as a viable branch of third-generation cognitive-behavioral therapies (Wilson et al. 2019). Like many other researchers, we became interested in self-compassion as a possible *protective* factor against the development of mental health problems and conducted a number of empirical studies on the topic. From the beginning of our research efforts, we were critical about the proper assessment

of this construct by means of the Self-Compassion Scale (SCS) (Neff 2003a). We noted that half of its items were fused with symptoms of psychopathology, and so we decided to remove them from the scale. However, during the peer review of our first paper (Muris et al. 2016a), we started to discover that a critical view on self-compassion and its scale was embraced with a skeptic eye. This increased our skeptical scientific attitude regarding the SCS: we did not understand why researchers did not see the obvious point that we were making regarding this measure, and hence, we began to share our critical thoughts with the scientific community (Muris 2016; Muris and Petrocchi 2017; Muris et al. 2016b, 2019b). However, the developer of the SCS (Neff 2016a, b, 2019) maintained that the scale in its original form provides a good index of self-compassion.

In the present article, we do not want to reiterate the entirety of the arguments that have been brought forward in these publications but rather provide a description of the main controversy and use this as an illustration for the process (and lack of progress) of science. We are specifically focusing on personal, cognitive, and social mechanisms at work preventing the correction and adjustment of theoretical notions and applied assessment instruments. Following this, we highlight new advancements in the field of self-compassion and come with clear recommendations regarding the assessment of this possibly protective construct.

Neff’s Theory and Assessment of Self-Compassion

A healthy attitude to oneself is considered to be one of the linchpins of resilience and the preservation of mental health (Baumeister and Vohs 2004). For a long time, the research on the role of self-related characteristics in people’s psychological functioning was predominantly focused on self-esteem, which refers to a person’s subjective evaluation of his/her own worth (Rosenberg 1965). Although research has shown that high self-esteem may buffer against mental health problems while low levels of self-esteem might increase the risk for the development of such difficulties, it has also been noted that self-esteem is quite resistant to change (e.g., Josephs et al. 2003) and thus seems to be a less suitable target for intervention. Around the beginning of this century, Western psychology started to exhibit interest in self-compassion, an alternative positive self-related concept originating from Buddhist culture. Pioneering empirical work was conducted by Neff (2003b) who defined self-compassion as “being touched by and open to one’s own suffering, not avoiding or disconnecting from it, generating the desire to alleviate one’s suffering and to heal oneself with kindness. [It] also involves offering nonjudgmental understanding to one’s pain, inadequacies and failures, so that one’s experience is seen as part

of the larger human experience” (p. 87). In a more recent elaboration of this definition, Neff conceptualized self-compassion as a “balance between increased compassionate and decreased uncompassionate self-responding to personal struggle” (Neff et al. 2018b, p. 371), which involves three key elements on bipolar ends: (1) being kind and supportive to oneself rather than harsh and judgmental, (2) recognizing that such difficulties constitute a normal part of humans’ life rather than feeling isolated from other people as a result of one’s imperfection, and (3) keeping the personal suffering in rational awareness rather than becoming fully absorbed by one’s problems (see Neff 2003a, b).

In line with this conceptualization, Neff (2003a) developed the SCS, a 26-item self-report questionnaire for measuring individual differences in “the three main components of self-compassion on separate subscales (self-kindness versus self-judgment, common humanity versus isolation, and mindfulness versus over-identification) with the intention of summing the subscale scores to create a total score that would represent a participants’ overall level of self-compassion” (p. 226). Table 1 shows examples of SCS items on the three dimensions, each contrasting compassionate (positive) and uncompassionate (negative) ways of self-responding. A shortened version of the scale (the SCS-SF) has also been created (Raes et al. 2011), which includes the 12 best loading items, 4 for each dimension with again an equal number for compassionate and uncompassionate self-directed responses.

Solid evidence exists for the basic psychometric properties of the SCS. The initial evaluation of the test by Neff (2003a) showed that the reliability of the scale was good, and this was true for both the internal consistency (Cronbach’s alpha was .92 for the total score and ranged between .75 and .81 for various subscales) and test-retest stability (test-retest correlations over a 3-week period were .93 for the total score and

between .80 and .88 for the subscales). Furthermore, support was found for the validity of the scale. Specifically, the total SCS score was positively related to scores on other positive self-related traits (i.e., self-esteem, self-acceptance, self-determination) while it was negatively associated with symptom levels of anxiety and depression, which is of course in line with the hypothesized protective qualities of self-compassion. In addition, it was found that Buddhists—who practice a type of meditation enhancing mindfulness and compassion—displayed, as expected, a higher SCS score than a comparison group of undergraduate students.

The only aberrant finding in the Neff (2003a) study was that no support could be documented for the dimensional nature of the three self-compassion components. Instead, confirmatory factor analysis showed that for each component, a two-factor solution provided a better fit for the data than the one-factor solution, which made Neff conclude that six separate but correlated factors in the SCS existed. In her view, these factors jointly constitute the overarching construct of self-compassion, which as such still justifies the employment of the total score.

The SCS has become highly popular in the academic field. The SCS was translated and validated in at least 17 countries. In most cases, the psychometric qualities of the scale have been reported to be just as favorable as those of the original English version. This is also true for the SCS-SF: the shortened scale is reliable and has a near-perfect correlation with the original 26-item scale (Raes et al. 2011) suggesting that the use of this economic measure is likely to produce the same results as obtained with its full-length counterpart.

Flourishing Research on Self-Compassion

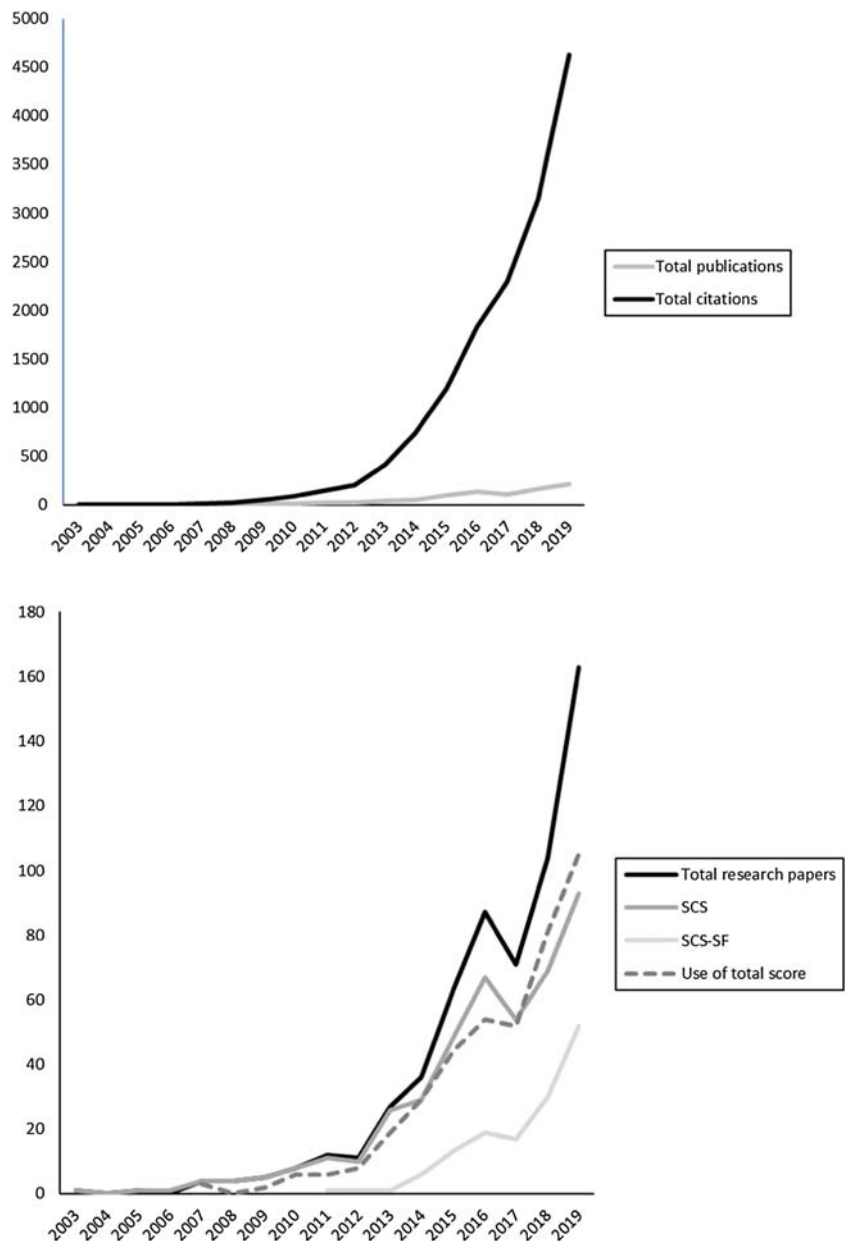
Since its introduction in the scientific literature, the construct of self-compassion has garnered considerable attention of the research community. A search of the literature conducted on December 31, 2019 in the Web of Science database using [SELF-COMPASSION in title] as the search term yielded 927 publications, of which 597 were empirical studies. Most of these investigations ($n = 571$, 95.6%) used the SCS or the SCS-SF (which became only available in 2011) to measure self-compassion, showing that these are the dominant instruments in the field. Alternative scales such as the Fear of Compassion for the Self (FCS) (Gilbert et al. 2011), Self-Other Four Immeasurables (SOFI) (Kraus and Sears 2009), and the Self-Compassion and Self-Criticism Scales (SCCS) (Falconer et al. 2015) are available but are less frequently used.

As can be seen in the top panel of Fig. 1, the number of publications has continuously increased over the years, and this also counts for the number of citations which is growing exponentially. The full-length version of the SCS is still most

Table 1 Item examples for each of the three self-compassion dimensions of Neff’s (2003a) Self-Compassion Scale (SCS)

SCS dimensions	Compassionate self-responding	Uncompassionate self-responding
1. Self-kindness ^a versus self-judgment ^b	^a I’m kind to myself when I’m experiencing suffering	^b When I see aspects of myself that I do not like, I get down on myself
2. Common humanity ^c versus isolation ^d	^c When I feel inadequate in some way, I try to remind myself that feelings of inadequacy are shared by most people	^d When I fail at something that’s important to me I tend to feel alone in my failure
3. Mindfulness ^e versus over-identification ^f	^e When something upsets me, I try to keep my emotions in balance	^f When something painful happens, I tend to blow the incident out of proportion

Fig. 1 Total number of publications and citations on self-compassion per year (top panel) and number of annual research publications split by the method used to assess this individual difference variable in various studies (bottom panel). *SCS* Self-Compassion Scale, *SCS-SF* Self-Compassion Scale-Short Form



popular, although the percentage of studies employing the SCS-SF is steadily increasing (from 8.3% in 2011 to 31.9% in 2019). Most researchers (72.0%) use the total score of the SCS or SCS-SF to obtain an overall indicator of self-compassion, and only in a minority of studies (28.0%) other scores are derived from these measures, such as scores on individual subscales or combined positive and negative subscales, and this modus operandi does not seem to change over time (bottom panel of Fig. 1). The trend to employ the total score is more prominent with the SCS-SF (89.8%) than with the SCS (65.1%), which is not surprising given the limited number of items included in the form which precludes the measurement of reliable factor/subscale scores.

The attraction of self-compassion lies in the fact that this construct may be relevant for understanding people's adjustment to life adversity and personal problems (Neff 2003b). It is clearly advocated as a protective mechanism, as we read in a recent writing by Neff and Germer (2017): "Self-compassion is a powerful way to enhance intrapersonal and interpersonal well-being. When we are mindful of our suffering and respond to it with kindness, remembering that suffering is part of the shared human condition, it appears that we are able to better cope with life's struggles" (p.382). In other words, it is assumed that self-compassion promotes psychological resilience by enabling the person to use more adaptive coping and emotion regulation strategies (Inwood and Ferrari 2018),

and this could help the individual to maintain healthy psychological functioning and shield against the development of mental health problems. Many studies have focused on the protective function of self-compassion, and although most of them are cross-sectional (correlational) in nature (but see Donald et al. 2018; Kirschner et al. 2019), this research has generally indicated that this trait is positively related to indices reflecting personal well-being (e.g., happiness, life satisfaction; Zessin et al. 2015) and negatively associated with measures tapping symptoms of psychopathology (e.g., anxiety, depression, stress; MacBeth and Gumley 2012). Furthermore, while most studies have concentrated on the relevance of self-compassion for people's adaptation within a clinical psychology context, the trait is also increasingly investigated as a positive psychological characteristic in stressful work-related (e.g., burn-out, motivation, procrastination), sports, and medical (e.g., HIV, chronic pain, cancer) settings (see for an overview: <https://self-compassion.org/>).

An additional appealing feature of self-compassion is that it also has much potential for therapy and interventions: unlike self-esteem, this trait appears to be amenable to change (Germer and Neff 2019). Specific treatments have been developed with the purpose to bolster compassionate self-responding (Kirby et al. 2017; Wilson et al. 2019), and indeed, evidence has been obtained that suggests that these interventions not only successfully increase self-compassion but also promote personal well-being and reduce psychopathology. Not surprisingly, many clinicians are enthusiastic and view self-compassion as a potent vehicle along which psychological functioning can be improved and human suffering can be eliminated.

Critique on the SCS

At first sight, the research on self-compassion is in good shape: a vast amount of evidence seems to indicate that this trait acts as a protective individual difference variable that seems to have more explanatory power than other positive psychology constructs (e.g., self-esteem, e.g., Neff and Vonk 2009; mindfulness, e.g., Van Dam et al. 2011) and also provides a lead for (improving) treatment in clinical settings. A weak point, however, is that the main body of knowledge is solely based on the SCS and that questions have been raised regarding the validity of this instrument. We argue that this criticism is far from trivial and seriously undermines the scientific foundation of the self-compassion concept.

The main point of critique regarding the SCS concerns the inclusion of items referring to ways of uncompassionate self-responding. These items measuring self-judgment, isolation, and over-identification were initially included in the scale as reversed items of the three key components of self-kindness, common humanity, and mindfulness. However, as is often the

case with reversed items, they can form separate factors and unintentionally introduce new dimensions and/or unwanted method variance in a measure (e.g., Wong et al. 2003). A recent study has indeed shown that this also applies to the SCS (Montero-Marin et al. 2018). The inclusion of the uncompassionate self-responding items is especially problematic because they are clearly in contrast with the protective nature of the self-compassion construct. Face validity checks performed on SCS items have revealed that whereas the compassionate self-responding components are regarded as aspects of positive cognitive coping and healthy psychological functioning, the uncompassionate self-responding components are mainly seen as indicators of vulnerability, psychopathological symptoms, and mental illness (Muris et al. 2018; see also Fig. 2). This result has also been confirmed by empirical data showing that while the compassionate SCS components of self-kindness, common humanity, and mindfulness are connected to adaptive personality features (e.g., optimism), positive mood states (e.g., happiness), and aspects of well-being (e.g., quality of life), their uncompassionate counterparts of self-judgment, isolation, and over-identification are more clearly linked to negative affect and symptoms of anxiety and depression (e.g., Neff et al. 2018a). Thus, the SCS contains multiple dimensions that are differentially related to external constructs (see also Brenner et al. 2017, 2018; Lopez et al. 2015, 2016; Muris et al. 2018, 2019a). It is clear that this has important consequences when researchers use and only report the total score of such a measure as one cannot know the nature of the different dimensions' contributions to that score as well as their unique predictive value for external variables (Smith et al. 2009).

Within the context of psychopathology, the inclusion of uncompassionate self-responding items in the SCS total score is questionable for another and perhaps more important reason. Several authors have argued that the components of self-judgment, isolation, and over-identification reflect a number of toxic processes that can be commonly observed in a wide range of mental health problems, especially those of an internalizing nature. Specifically, self-judgment shows clear similarity with harsh self-criticism and self-punishment, isolation shares features with social withdrawal and loneliness, whereas over-identification matches with self-absorption and self-focused rumination (Körner et al. 2015; Lopez et al. 2018; Muris 2016). Whether the uncompassionate self-responding components of the SCS reflect some underlying vulnerability factor (e.g., neuroticism; Pfattheicher et al. 2017) or directly reflect symptoms of psychopathology (Muris et al. 2016b) is still a matter of debate, but the fact is that there is (too) little awareness in the research on self-compassion that this type of "measurement confounding" (see, e.g., Lemery et al. 2002) occurs, in particular when using the measure as a "predictor" of mental health problems. As noted earlier, the vast majority of researchers in the field of self-compassion appears to use

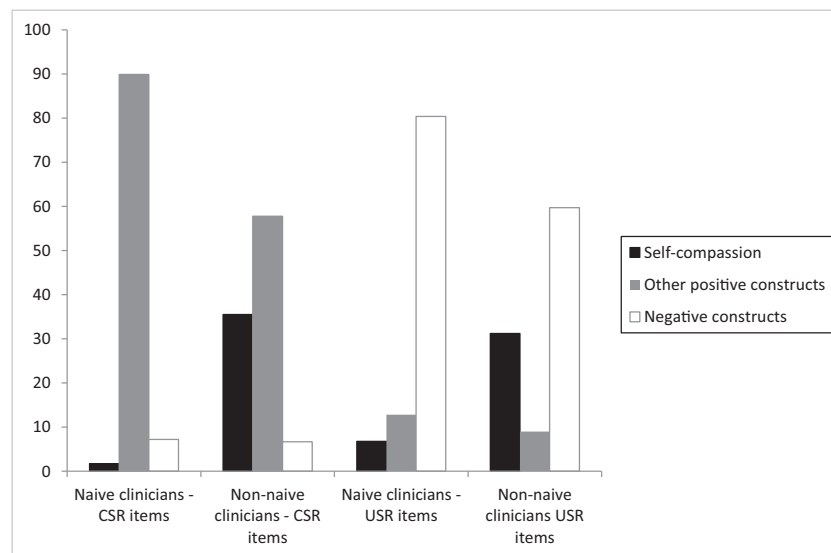


Fig. 2 New data on the face validity of the SCS (Corsius 2018): Percentages of naïve and non-naïve clinicians assigning compassionate (CSR) and uncompassionate self-responding (USR) items of the SCS to the categories of self-compassion, other positive features, and negative features. Results showed that compassionate self-responding items were more often classified as positive constructs whereas uncompassionate self-responding items were more frequently identified as negative constructs, and this appeared true for naïve as well as non-naïve clinicians.

Further, as expected, the non-naïve clinicians were better in classifying self-compassion items than the naïve clinicians, although it should also be noted that even their scores were not impressive: that is, 35.5% of the compassionate and 31.2% of uncompassionate self-responding items were correctly identified as belonging to the self-compassion construct. SCS Self-Compassion Scale. Naïve clinicians—clinicians who were not familiar with the self-compassion construct, non-naïve clinicians—clinicians who were familiar with the self-compassion construct

the total SCS score (which includes the reversed uncompassionate self-responding items) and without further consideration treat this as a valid indicator of a protective trait.

A Scientific Smoke Curtain to Defend the Use of the SCS Total Score

The most important argument put forward to defend the use of the SCS total score is based on the factor analytic finding that there seems to be a common overarching factor that serves as an umbrella for the six compassionate and uncompassionate components included in the measure (Neff 2016b, 2019). An exemplary study was conducted by Neff et al. (2019) who used sophisticated statistical methodology to explore the factor structure of the SCS using the data of 11,685 participants from 20 diverse samples. The analysis revealed that a single bi-factor model, in which each SCS item not only loaded on its corresponding factor but also on an overarching factor, provided the best fit for the data, thereby justifying the use of the total score (see also Cleare et al. 2018; Neff et al. 2018b; Toth-Kiraly et al. 2017).

However, there has been more research investigating the factor structure of the SCS, and the outcomes have been mixed: quite a number of studies obtained support for a correlated six-factor model (Benda and Reichova 2016; Castilho et al. 2015; Cunha et al. 2016; Garcia-Campayo et al. 2014; Hupfeld and Ruffieux 2011; Kotsou and Leys 2016;

Kumlander et al. 2018; Petrocchi et al. 2014; Pfattheicher et al. 2017; Ursic et al. 2019) while there are also other investigations pointing in the direction of a solution with two overarching factors representing compassionate and uncompassionate self-responding (Brenner et al. 2017; Coroiu et al. 2018; Costa et al. 2016; Halamova et al. 2018; Hayes et al. 2016; Zhang et al. 2019). Taken together, it can be concluded 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. The structure of the scale is complex, and the compassionate and uncompassionate self-responding components need to be split while at the same time allowing them in some way to share variance (Williams et al. 2014).

Examination of the factor structure of a scale is certainly important as it enables us to determine whether the instrument adequately captures the various components that constitute a theoretical construct. However, it is important to note that this type of analysis only covers one aspect of validity, namely construct validity, and that it should not be used as a scientific smoke curtain to obscure the fact that there are other problems with a measure. The divergent correlations between the compassionate and uncompassionate self-responding components of the SCS and external measures, and the confounding of the uncompassionate components with vulnerability and psychopathology, simply cannot be ignored.

Implications of Using the SCS Total Score for Research on Self-Compassion

We signal at least four problems for the scientific inquiry of self-compassion with this sole focus on the SCS total score. First of all, the true protective value of self-compassion is obscured. On the basis of meta-analytic evidence, Muris and Petrocchi (2017) concluded that the uncompassionate self-responding components of the SCS are more strongly related to mental health problems than the compassionate self-responding components, implying that when using the total score, the link with psychopathology will be inflated. Direct support for this idea is still sparse, but in a recent paper, we demonstrated that when statistically correcting for uncompassionate self-responding, the contribution of compassionate self-responding to symptoms of anxiety, depression, stress, and related constructs was significantly reduced, explaining only a marginal proportion in the variance of these psychopathology indicators (Muris et al. 2019b; see Fig. 3).

A second problem is that the inclusion of the uncompassionate self-responding components in the SCS total score will hinder researchers from investigating the precise role of self-compassion in the multifactorial origins of psychopathology and from making a fair comparison with other predictor variables. For example, in one of our own studies (Muris et al. 2016a), we examined the value of self-compassion as a protective variable against symptoms of anxiety and depression as compared to other positive self-related factors such as self-esteem and self-efficacy. We used a SCS version from which

the uncompassionate self-responding items had been discarded; if we had not done so, this would probably have resulted in the finding that self-compassion had incremental validity in predicting psychopathological symptoms over the other positive factors (cf. Neff and Vonk 2009), which now did not turn out to be the case (i.e., self-esteem and self-efficacy—and not self-compassion—were the variables found to have unique explanatory power). Thus, the employment of the SCS total score will obscure to what extent a protective trait like self-compassion really stands the competition with other etiological factors of mental health problems.

A third problem with the employment of the total score is that it may lead researchers to delve into issues that are perhaps trivial. A good example is the investigation of gender differences in self-compassion. A meta-analysis by Yarnell et al. (2015) has indicated that—when considering the SCS total score—women are in general less self-compassionate than men, on the basis of which it has been recommended to explore in future studies whether proper training may help women foster more compassion toward the self and make them more resilient to the development of mental health problems. In a subsequent study, Yarnell et al. (2019) examined whether the gender difference in self-compassion can be explained by differences in gender role orientation. The results were quite complex, but the overall picture suggests that “masculinity was the most consistent positive predictor of self-compassion” (p.1147, meaning that in both men and women, participants with higher levels of masculine traits displayed higher total SCS scores) while the role of femininity

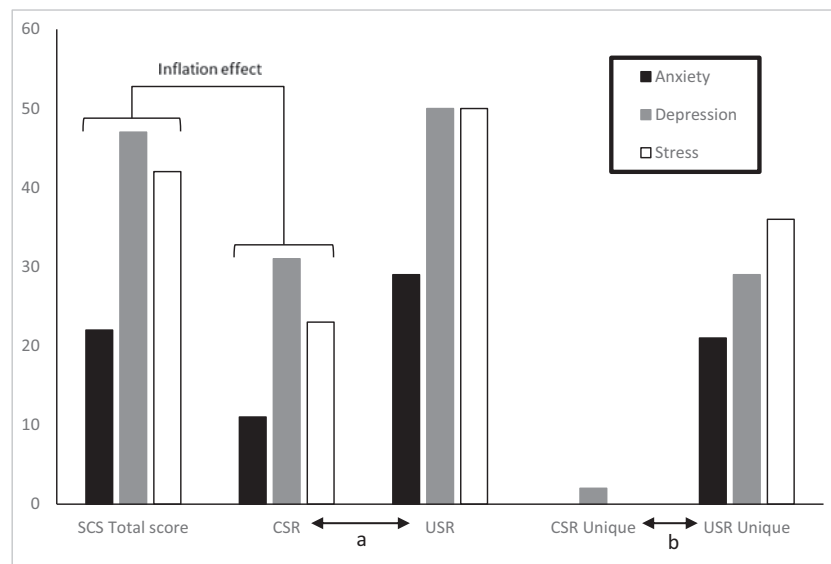


Fig. 3 The relation between SCS scores and some key indices of psychopathology expressed in percentages of shared variance. As can be seen, when using the SCS total score, there appears to be a robust relation between self-compassion and symptoms of anxiety, depression, and stress. Note, however, that (a) the shared variance between uncompassionate self-responding (USR) and symptoms is about twice as large as the shared variance between compassionate self-responding (CSR) and

symptoms (all comparisons are significant at $p < .001$) and (b) when controlling for the overlap between USR and CSR, only USR shows a unique link to symptoms. On the basis of these findings, it is hard to evade the conclusion that the incorporation of USR in the SCS inflates the relation between self-compassion and symptoms of psychopathology (Muris et al. 2019b)

was less clear. To explain the intricate findings, Yarnell et al. (2019) point at the multiplistic nature of gender roles and the way these are assessed, but they forget to note that self-compassion as measured with the SCS also has multiple faces. Studies that did explore gender differences for SCS subscales indicated that women only scored higher on self-judgment, isolation, and over-identification (e.g., Bluth and Blanton 2015), implying that their lower total self-compassion scores do not truly reflect lower levels of compassionate self-responding but are mainly due to their higher endorsement of uncompassionate self-responding items. The latter is obviously less groundbreaking as it is already well known in the literature that women generally score higher on negative, internalizing psychopathology-related features than men (e.g., Kramer et al. 2008).

A fourth and final problem is that the use of the SCS total score will conceal the precise and unique role of self-compassion during psychological treatment. As noted earlier, an important assumption is that self-compassion is a modifiable characteristic, and in the past years, various types of interventions have been developed to examine whether these might reduce psychopathology (Kirby et al. 2017; Wilson et al. 2019). However, as rightly noted by Kirby and Gilbert (2019), by lumping together compassionate and uncompassionate self-responding into one score, it will remain unclear what is actually happening during therapy. Obviously, the abolishment of uncompassionate self-responding is of less interest: given its commonality with psychopathology, this seems to merely reflect symptom reduction, which is of course an important target of almost any therapeutic intervention. It seems more crucial and interesting to focus on the change in self-compassionate responding and to study whether this represents the mechanism responsible for the observed treatment effects (Wadsworth et al. 2018).

Altogether, it is inevitable to conclude that, especially within a context of mental health problems and stress, the inclusion of the uncompassionate self-responding components in the SCS is problematic. And although we and several other scholars (Brenner et al. 2018; Kirby and Gilbert 2019; Lopez et al. 2018) have repeatedly emphasized this issue, most researchers appear to show little appreciation for these critical notes and continue to use the SCS total score including the (reversely scored) uncompassionate self-responding items as the preferred index of self-compassion.

The Process of Science

Why do the critical points raised regarding the SCS and the empirical data collected to substantiate this criticism not lead to an adjustment in the theory of self-compassion and more specifically an altered employment of the scale in research? Or in other words: why is science apparently not operating as a

Popperian process? The short answer is that this is because 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.

First of all, it is not uncommon that scientists continue to defend a theory against contrasting evidence and arguments. A famous example was Charles Darwin, a brilliant researcher who early in his career as a geologist was puzzled by the “parallel roads” of Glen Roy, three perfectly horizontal terraces along the mountainsides in the Lochaber area of the Scottish Highlands (Rudwick 1974). Darwin advanced the theory that the “parallel roads” were raised beaches of marine origin, and although he could not find any evidence in support of this account (e.g., fossils of marine animals and plants), he vigorously defended it, in spite of the fact that a more plausible theory had been formulated stating that the terraces in the landscape were cut by the waxing and waning shores of a glacier lake during the ice age. Only short before his death, Darwin recognized that he had it all wrong and admitted that his account of the “parallel roads” of Glen Roy was “one long gigantic blunder”.

Interestingly, Chinn and Brewer (1993, 1998, 2001) have described a framework describing how researchers deal with data that are not in keeping or even in contrast with their theory. In this framework, eight possible ways are specified of how scientists respond to such anomalous data. The most common responses reflect defensive maneuvers such as ignoring, rejection, exclusion, abeyance, and reinterpretation, all of which have the net effect that the initial theory remains unchanged. There are a number of reasons why scientists choose to preserve a theory in spite of anomalous data (Chinn and Brewer 1993). The first reason has to do with a person’s beliefs with regard to the theory. If a theoretical belief is deeply entrenched and associated with a strong personal commitment, it will be challenging to change this notion. Many researchers in the self-compassion research field are scientist-practitioners who operate in science because they are invested in developing diagnostic tools and treatment methods with the ultimate goal of helping patients more effectively in clinical practice. However, researchers with a scientist-practitioner focus often show less concern regarding validity checks and other psychometric issues, although this putatively is a threat to science (Lilienfeld et al. 2015). Thus, researchers tend to “blindly” employ a measure simply because it is continuously advocated as the ideally suited index for measuring a construct not only by a leading researcher (authority bias—the tendency to attribute greater value to the opinion of an authority figure; Milgram 1963) but also by many coworkers in this research field (groupthink or bandwagon effect—the tendency to do or believe things because many other people do or believe the same; Mailoo 2015).

Another ground for why aberrant data tend to be neglected is concerned with the lack of a solid alternative theory.

Meanwhile, it is important to note that Gilbert, who is another well-respected researcher in the self-compassion field, has formulated an alternative theoretical account that might be able to effectively deal with the main problem of the SCS. This account, known as the theory of social mentalities (Gilbert 2000, 2005), proposes that there are three separate brain-based systems guiding people's social behavior of which two are of particular importance for the current discussion: (1) the (parasympathetic) safeness system that elicits thoughts, feelings, and actions promoting positive relationships with others and the self, which is thought to be involved in compassionate self-responding, and (2) a (sympathetic) threat-defense system that prompts thoughts, feelings, and actions that mainly serve to reduce threat, which is related to ways of uncompassionate self-responding. Thus, there is a solid alternative theory (but see Khoury 2019) proposing that compassionate and uncompassionate self-responding indeed reflect different processes that are moderated by different brain systems (for further reading, see Klimecki and Singer 2017), which—as noted earlier—implies that it is not appropriate to combine them in a single score of self-compassion.

A further reason for theory-preserving responses to anomalous data relates to the credibility of the anomalous data. 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. For instance, the resistance to modification is likely to be guided by cognitive biases such as anchoring (i.e., relying too heavily on the first piece of information on a subject; Epley and Gilovich 2006), congruence bias (i.e., the tendency to test a hypothesis exclusively in one particular way instead of testing possible alternative hypotheses; Iverson et al. 2008), and confirmation bias (i.e., the tendency to search for, interpret, focus on, and remember information in a way that confirms one's preconceptions; Oswald and Grosjean 2004). In this light, it is also important to note that the SCS has been the cornerstone of the self-compassion literature, and so, an acknowledgement of the questionable validity of the scale would lead to a re-analysis and reinterpretation of most of the collected data (with an unknown, probably less advantageous outcome for the self-compassion construct), which of course feeds cognitive distortions such as conservatism (i.e., the tendency to revise one's belief insufficiently when presented with new

evidence; Iverson et al. 2008), the Ostrich effect (i.e., ignoring an obvious negative situation; Karlsson et al. 2009), and status quo bias (i.e., the tendency to like things to stay the same; Nebel 2015).

Most of the abovementioned cognitive biases operate at an automatic level, which means that scientists make these types of information processing errors without being (fully) aware of them. However, 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 (see Tijdink et al. 2016). In the case of self-compassion, it is more profitable to employ the SCS total score and treat the construct as a pure protective factor rather than to present a more nuanced (although more honest) picture including the role of compassionate and uncompassionate self-responding. Obviously, the second option carries the risk that results will show that, in the context of psychopathology, the uncompassionate ways of self-responding have more predictive power than their compassionate counterparts, a finding that has less scientific news value. There are examples of researchers in the self-compassion field who in spite of their apparent awareness of the potential problem with the SCS (as they previously published a paper displaying separate results for the compassionate and uncompassionate self-responding components) completely neglect this point when writing their next research report. This comes close to a phenomenon that has been labeled as “cherry-picking” (Murphy and Aguinis 2019).

Peer review is another element in science that may be associated with various types of biases serving to preserve an existing theoretical account and hinder the assimilation of new points of view (Haffar et al. 2019). As noted in the introductory section of this paper, reviewers are also prone to processes that distort their judgment when evaluating the empirical papers of researchers (who are often colleagues and scientific peers). Here too, anchoring, confirmation bias, and conservatism are at work in conjunction with more specific distortions such as the availability cascade (i.e., a notion that is increasingly repeated in the scientific literature will also be considered as more plausible; Kuran and Sunstein 1999). In the case of the SCS, the law-of-the-instrument effect may also be operating, which refers to an over-reliance on a familiar tool and its prescribed scoring method (Kaplan 1964). Journal editors have a decisive role in the peer review process but are of course susceptible to the same set of biases as the reviewers. They should be aware of these distortions, try to make objective judgments, and make the effort to encourage critical scientific dialog, which will ultimately lead to an advancement in knowledge.

New Insights in Self-Compassion

It is important to note that Neff's definition of self-compassion—on the basis of which she developed the SCS—was based on her own observations and Buddhist readings. There is nothing wrong to adopt such an approach, and her conceptualization has proven to be powerful enough to attract attention from other researchers and clinicians. In the meantime, theoretical notions on compassion in general and self-compassion in specific have considerably advanced during the past years (Seppälä et al. 2018). After an extensive review of the literature on the various ways that (self-)compassion has been conceptualized, Strauss et al. (2016) came to the conclusion that the construct essentially consists of five elements: (1) recognizing suffering, (2) understanding the universality of human suffering, (3) feeling for the person suffering, (4) tolerating uncomfortable feelings, and (5) acting or motivation to act to alleviate suffering. Acknowledging the fact that most existing measures do not comprehensively measure self-compassion, these researchers developed a new questionnaire, the Sussex-Oxford Compassion Scales (SOCS), of which the 20-item self-related version measures self-compassion. In a first psychometric test (Gu et al. 2020), the scale was demonstrated to possess good reliability and validity. In support of its concurrent validity, scores on the new measure were found to correlate positively with the total score of the short form of the SCS: the SOCS subscales referring to feeling for the person suffering, tolerating uncomfortable feelings, and acting or motivation to act to alleviate suffering showed the most substantial correlations ($>.50$) with the SCS, indicating that these aspects are reasonably well represented by Neff's operationalization of self-compassion.

A post hoc analysis performed on these data investigating the unique relations between compassionate and uncompassionate self-responding components of the SCS and various

SOCS subscales revealed that compassionate self-responding was a significantly better predictor of various SOCS self-compassion elements than uncompassionate self-responding (Fig. 4). This suggests once more that if one wants to assess the true nature of self-compassion by means of the SCS, it is preferable to rely on the compassionate self-responding components. A similar conclusion was reached by Montero-Marín et al. (2018) who adopted a multitrait-multimethod analytical procedure analyzing the SCS data from 4120 participants in 11 samples with different cultural backgrounds. Their analysis indicated that “the positively valenced [compassionate self-responding] items, compared with the negative [uncompassionate self-responding] ones (which suffered more from method effects), were better explained by the corresponding trait factors of self-compassion” (p. 10), which led them to propose that the assessment of self-compassion with the SCS can best be confined to the compassionate items.

Concluding Remarks

Science involves the observation, identification, description, experimental investigation, and theoretical explanation of phenomena occurring in the real world. Via publications in peer-reviewed journals, scientists continuously inform the world about new insights and advancements they have made. According to Popper (1963), if the new data presented in a paper are in line with a theory, that theory is confirmed and consolidated. However, when the data are not in agreement with a theory, that should result in either a modification or rejection of the theory. In Popper's view, science operates as an objective rational process, but the reality is that this is certainly not always the case. Science is a socially driven enterprise conducted by human beings who—especially when

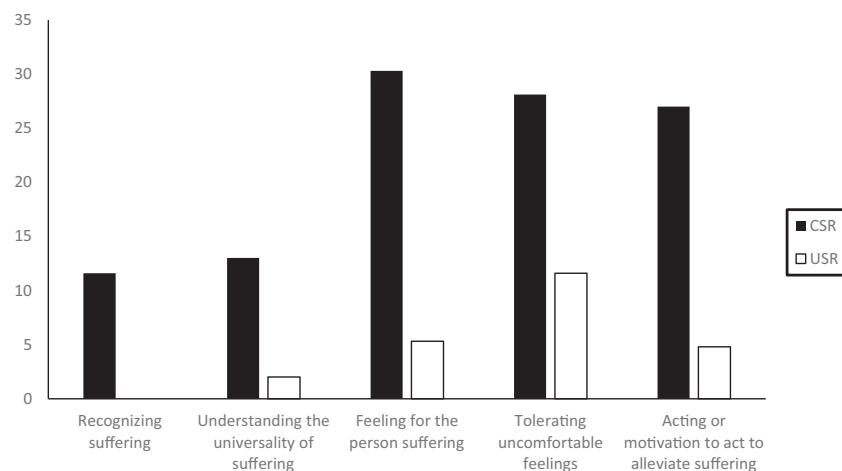


Fig. 4 The relation between compassionate and uncompassionate self-responding (CSR and USR) scores of the SCS and elements of self-compassion as measured by the Sussex-Oxford Compassion Scale (SOCS) as expressed in percentages of shared variance (while controlling

for the overlap between CSR and USR). The results indicate that CSR is a statistically significantly better indicator of self-compassion elements than USR (all p 's $< .001$). Analysis conducted on data of the health care staff sample ($N = 1158$) in the Gu et al. (2020) study

dealing with personally relevant topics—are prone to all kinds of biases. As a result, progress in science is not occurring readily and automatically. It really takes time for a research field to let go of a firmly established theory because of non-fitting evidence.

We have taken the research field of self-compassion as an illustration of the true process of science. 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 and not in keeping with the true nature of this positive psychology construct (see also Khoury 2019). Nonetheless, the researchers in this domain do not seem ready for what Kuhn (1962) called a “paradigm shift”. Should it occur, such a shift would imply a theoretical change as well as a change in the measurement of self-compassion. The work by Strauss et al. (2016) has led to a more comprehensive description of the construct, on the basis of which a new scale (the SOCS; Gu et al. 2020) has been developed. This measure reliably assesses various aspects of self-compassion, the relevance of which can be further investigated in future studies. In the case of a continued use of the SCS, the solution is obvious: researchers should only use the compassionate self-responding components of the scale in case they are interested in studying the protective nature of the construct or consistently make a distinction between compassionate and uncompassionate ways of self-responding to study their unique and interactive effects within the context of psychopathology, thereby treating them for what they really are: separate concepts of protection and vulnerability.

Acknowledgements The authors kindly thank Dr. Clara Strauss and Dr. Jenny Gu (University of Sussex) for conducting the post hoc analysis on the relation between the Sussex-Oxford Compassion Scale and the Self-Compassion Scale.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Approval All procedures were in accordance with ethical standards. The research protocol was approved by the Ethical Review Committee Psychology and Neuroscience at Maastricht University, The Netherlands.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Barrett, L. F. (2017). *How emotions are made: the secret life of the brain*. Boston: Houghton Mifflin Harcourt.
- Baumeister, R. F., & Vohs, K. D. (2004). *Handbook of self-regulation. Research, theory, and applications*. New York: The Guilford press.
- Bayne, T., Brainard, D., Byrne, R. W., Chittka, L., Clayton, N., Heyes, C., Mather, J., Ölveczky, B., Shadlen, M., Suddendorf, T., & Webb, B. (2019). What is cognition? *Current Biology*, 29, R608–R615.
- Benda, J., & Reichova, A. (2016). Psychometrické charakteristiky Ceske verze Self-Compassion Scale (SCS-CZ). *Ceskoslovenska Psychologie*, 60, 120–136.
- Bluth, K., & Blanton, P. W. (2015). The influence of self-compassion on emotional well-being among early and older adolescent males and females. *Journal of Positive Psychology*, 10, 219–230.
- Brenner, R. E., Heath, P. J., Vogel, D. L., & Crede, 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.
- 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, 346–357.
- Castilho, P., Pinto-Gouveia, J., & Duarte, J. (2015). Evaluating the multifactor structure of the long and short versions of the Self-Compassion Scale in a clinical sample. *Journal of Clinical Psychology*, 71, 856–870.
- Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: a theoretical framework and implications for science instruction. *Review of Educational Research*, 63, 1–49.
- Chinn, C. A., & Brewer, W. F. (1998). An empirical test of a taxonomy of responses to anomalous data in science. *Journal of Research in Science Teaching*, 35, 623–654.
- Chinn, C. A., & Brewer, W. F. (2001). Models of data: a theory of how people evaluate data. *Cognition and Instruction*, 19, 323–393.
- 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.
- Coroiu, A., Kwakkenbos, L., Moran, C., Thombs, B., Albani, C., Bourkas, S., Zenger, M., Brahler, E., & Körner, A. (2018). Structural validation of the Self-Compassion Scale with a German general population sample. *PLoS One*, 13, e0190771.
- Corsius, N. (2018). *Further exploration of the face validity of the Self-Compassion Scale* (unpublished master's thesis). Maastricht University, Maastricht, The Netherlands.
- Costa, J., Maroco, 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 and Psychotherapy*, 23, 460–468.
- 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.
- Donald, J. N., Ciarrochi, J., Parker, P. D., Sahdra, B. K., Marshall, S. L., & Guo, J. (2018). A worthy self is a caring self: examining the developmental relations between self-esteem and self-compassion in adolescents. *Journal of Personality*, 86, 619–630.
- Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic. *Psychological Science*, 17, 311–318.
- Falconer, C. J., King, J. A., & Brewin, C. R. (2015). Demonstrating mood repair with a situation-based measure of self-compassion and self-criticism. *Psychology and Psychotherapy: Theory Research and Practice*, 88, 351–365.

- Garcia-Campayo, J., Navarro-Gil, M., Andres, E., Montero-Marin, J., Lopez-Artal, L., & Piva 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.
- Germer, C., & Neff, K. D. (2017). *Teaching the mindful self-compassion program. A guide for professionals*. New York: The Guilford Press.
- 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., McEwan, K., Matos, M., & Rivis, A. (2011). Fears of compassion: development of three self-report scales. *Psychology and Psychotherapy: Theory Research and Practice*, 84, 239–255.
- Gu, J., Baer, R., Cavanagh, K., Kuyken, W., & Strauss, C. (2020). Development and psychometric properties of the Sussex-Oxford Compassion Scales (SOCS). *Assessment*, 27, 3–20.
- Haffar, S., Bazerbachi, F., & Murad, M. H. (2019). Peer review bias: a critical review. *Mayo Clinic Proceedings*, 4, 670–676.
- Halamova, J., Kanovsky, M., & Pacuchova, M. (2018). Self-Compassion Scale: IRT psychometric analysis, validation, and factor structure—Slovak translation. *Psychologica Belgica*, 57, 190–209.
- Hayes, J. A., Lockard, A. J., Janis, R. A., & Locke, B. D. (2016). Construct validity of the Self-Compassion Scale-Short Form among psychotherapy clients. *Counselling Psychology Quarterly*, 29, 405–422.
- Hornsveld, H. (1996). *Farewell to the hyperventilation syndrome*. Doctoral thesis, University of Amsterdam, The Netherlands.
- Hupfeld, J., & Ruffieux, N. (2011). Validierung einer Deutschen version der Self-Compassion Scale (SCS-D). *Zeitschrift für Klinische Psychologie und Psychotherapie*, 40, 115–123.
- Inwood, E., & Ferrari, M. (2018). Mechanisms of change in the relationship between self-compassion, emotion regulation, and mental health. *Applied Psychology: Health and Well-Being*, 10, 215–235.
- Iverson, G. L., Brooks, B. L., & Holdnack, J. A. (2008). Misdiagnosis of cognitive impairment in forensic neuropsychology. In R. L. Heilbroner (Ed.), *Neuropsychology in the courtroom: expert analysis of reports and testimony* (pp. 243–266). New York: Guilford Press.
- Josephs, R. A., Bosson, J. K., & Jacobs, C. G. (2003). Self-esteem maintenance processes: why low self-esteem may be resistant to change. *Personality and Social Psychology Bulletin*, 29, 920–933.
- Kaplan, A. (1964). *The conduct of inquiry: methodology for behavioral science*. San Francisco: Chandler Publishing.
- Karlsson, M., Loewenstein, G., & Seppi, D. (2009). The ostrich effect: selective attention to information. *Journal of Risk and Uncertainty*, 38, 95–115.
- Khouri, B. (2019). Compassion: embodied and embedded. *Mindfulness*, 10, 2363–2374.
- Kirby, J. N., & Gilbert, P. (2019). Commentary regarding Wilson et al. (2018) "Effectiveness of 'self-compassion' related therapies: a systematic review and meta-analysis". All is not as it seems. *Mindfulness*, 10, 1006–1016.
- Kirby, J. N., Tellegen, C. L., & Steindl, S. R. (2017). A meta-analysis of compassion-based interventions: current state of knowledge and future directions. *Behavior Therapy*, 48, 778–792.
- Kirschner, H., Kuyken, W., Wright, K., Roberts, H., Brejcha, C., & Karl, A. (2019). Soothing your heart and feeling connected: a new experimental paradigm to study the benefits of self-compassion. *Clinical Psychological Science*, 7, 545–565.
- Klimecki, O. M., & Singer, T. (2017). The compassionate brain. In E. M. Seppälä, E. Simon-Thomas, S. L. Brown, M. C. Worline, C. D. Cameron, & J. R. Doty (Eds.), *The Oxford handbook of compassion science* (pp. 109–120). New York: Oxford University Press.
- 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.
- 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.
- Kramer, M. D., Krueger, R. F., & Hicks, B. M. (2008). The role of internalizing and externalizing liability factors in accounting for gender differences in the prevalence of common psychopathological syndromes. *Psychological Medicine*, 38, 51–61.
- Kraus, S., & Sears, S. (2009). Measuring the immeasurables: development and initial validation of the Self-Other Four Immeasurables (SOFI) scale based on Buddhist teachings on loving kindness, compassion, joy, and equanimity. *Social Indicators Research*, 92, 169–181.
- Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Kumlander, S., Lahtinen, O., Turunen, T., & Salmivalli, C. (2018). Two is more valid than one, but is six even better? The factor structure of the Self-Compassion Scale (SCS). *PLoS One*, 13, e0207706.
- Kuran, T., & Sunstein, C. R. (1999). Availability cascades and risk regulation. *Stanford Law Review*, 51, 683–768.
- Lakatos, I., & Musgrave, A. (1974). *Criticism and the growth of knowledge*. New York: Cambridge University Press.
- Lemery, K. S., Essex, M. J., & Smider, N. A. (2002). Revealing the relation between temperament and behavior problem symptoms by eliminating measurement confounding: expert ratings and factor analyses. *Child Development*, 73, 867–882.
- Ley, R. (1985). Blood, breath, and fears: a hyperventilation theory of panic attacks and agoraphobia. *Clinical Psychology Review*, 5, 271–285.
- Lilienfeld, S. O., Ritschel, L. A., Lynn, S. J., Cautin, R. L., & Latzman, R. D. (2015). Science-practice gap. In R. L. Cautin & S. O. Lilienfeld (Eds.), *The encyclopedia of clinical psychology* (pp. 2548–2555). Hoboken: Wiley and Sons.
- Lopez, 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.
- Lopez, A., Sanderman, R., & Schroevers, M. J. (2016). Mindfulness and self-compassion as unique and common predictors of affect in the general population. *Mindfulness*, 7, 1289–1296.
- Lopez, A., Sanderman, R., & Schroevers, M. J. (2018). A close examination of the relationship between self-compassion and depressive symptoms. *Mindfulness*, 9, 1470–1478.
- MacBeth, A., & Gumley, A. (2012). Exploring compassion: a meta-analysis on the association between self-compassion and psychopathology. *Clinical Psychology Review*, 32, 545–552.
- Mailoo, V. (2015). Common sense or cognitive bias and groupthink: does it belong in our clinical reasoning? *British Journal of General Practice*, 65, 27.
- McComas, W. F. (1996). Ten myths of science: reexamining what we think about the nature of science. *School Science and Mathematics*, 96, 10–16.
- Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, 67, 371–378.
- Montero-Marin, J., Kuyken, W., Crane, C., Gu, J., Baer, R., Al-Awamleh, A. A., Akutsu, S., Araya-Veliz, C., Ghorbani, N., Chen, Z. J., Kim, M. S., Mantzios, M., Rolim dos Santos, D. N., Serrano Lopez, L. C., Teleb, A. A., Watson, P. J., Yamaguchi, A., Yang, E., & Garcia-Campayo, J. (2018). Self-compassion and cultural values: a cross-cultural study of self-compassion using a multitrait-multimethod (MTMM) analytical procedure. *Frontiers in Psychology*, 9, 2638.

- 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.
- 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 and Psychotherapy*, 24, 373–383.
- Muris, P., Meesters, C., Pierik, A., & De Kock, B. (2016a). Good for the self: self-compassion and other self-related constructs in relation to symptoms of anxiety and depression in non-clinical youths. *Journal of Child and Family Studies*, 25, 607–617.
- Muris, P., Otgaar, H., & Petrocchi, N. (2016b). Protection as the mirror image of psychopathology: further critical notes on the Self-Compassion Scale. *Mindfulness*, 7, 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, 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, 196–199.
- Murphy, K. R., & Aguinis, H. (2019). HARKing: how badly can cherry-picking and question trolling produce bias in published results? *Journal of Business and Psychology*, 34, 1–17.
- Nebel, J. M. (2015). Status quo bias, rationality, and conservatism about value. *Ethics*, 125, 449–476.
- 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, 85–101.
- Neff, K. D. (2016a). Does self-compassion entail reduced self-judgment, isolation, and over-identification? A response to Muris, Otgaar, and Petrocchi (2016). *Mindfulness*, 7, 791–797.
- Neff, K. D. (2016b). The Self-Compassion Scale is a valid and theoretically coherent measure of self-compassion. *Mindfulness*, 7, 264–274.
- Neff, K. D. (2019). Setting the record straight about the Self-Compassion Scale. *Mindfulness*, 10, 200–202.
- Neff, K. D., & Germer, C. (2017). Self-compassion and psychological well-being. In E. M. Seppälä, E. Simon-Thomas, S. L. Brown, M. C. Worline, C. D. Cameron, & J. R. Doty (Eds.), *The Oxford handbook of compassion science* (pp. 371–385). New York: Oxford University Press.
- Neff, K. D., & Vonk, R. (2009). Self-compassion versus global self-worth: two different ways of relating to oneself. *Journal of Personality*, 77, 23–50.
- Neff, K. D., Long, P., Knox, M. C., Davidson, O., Kuchar, A., Costigan, A., Williamson, Z., Rohleder, N., Toth-Kiraly, I., & Breines, J. G. (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, 627–645.
- Neff, K. D., Toth-Kiraly, 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, 371–392.
- Neff, K. D., Toth-Kiraly, 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., Sirios, F. M., De Souza, L. K., Svendsen, J. L., Wilkinson, R. B., & Mantzios, M. (2019). Examining the factor structure of the Self-Compassion Scale in 20 diverse samples: support for use of a total score and six subscale scores. *Psychological Assessment*, 31, 27–45.
- Oswald, M. E., & Grosjean, S. (2004). Confirmation bias. In R. F. Pohl (Ed.), *Cognitive illusions: a handbook on fallacies and bias in thinking, judgment, and memory* (pp. 79–96). Hove: Psychology Press.
- Petrocchi, N., Ottaviani, C., & Couyoumdjian, A. (2014). 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.
- Popper, K. R. (1963). *Conjectures and refutations: the growth of scientific knowledge*. London: Routledge.
- Raes, F., Pommier, E., Neff, K. D., & Van Gucht, D. (2011). Construction and factorial validation of a short form of the Self-Compassion Scale. *Clinical Psychology and Psychotherapy*, 18, 250–255.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton: Princeton University Press.
- Rudwick, M. J. S. (1974). Darwin and Glen Roy: a 'great failure' in scientific method? *Studies in the History and Philosophy of Science*, 5, 97–185.
- Salkovskis, P. M., & Clark, D. M. (1990). Affective responses to hyperventilation: a test of the cognitive model of panic. *Behaviour Research and Therapy*, 28, 15–28.
- Schwartz, R. S., Lederman, N. G., & Crawford, B. A. (2004). Developing views of nature of science in an authentic context: an explicit approach to bridging the gap between nature of science and scientific inquiry. *Science Education*, 88, 610–645.
- Seppälä, E. M., Simon-Thomas, E., Brown, S. L., Worline, M. C., Cameron, C. D., & Doty, J. R. (Eds.). (2018). *The Oxford handbook of compassion science*. New York: Oxford University Press.
- Smith, G. T., McCarthy, D. M., & Zapolski, T. C. B. (2009). On the value of homogeneous constructs for construct validation, theory testing, and the description of psychopathology. *Psychological Assessment*, 21, 272–284.
- Strauss, C., Taylor, B. L., Gu, J., Kuyken, W., Baer, R., Jones, F., & Cavanagh, K. (2016). What is compassion and how can we measure it? A review of definitions and measures. *Clinical Psychology Review*, 47, 15–27.
- Tijdkink, J. K., Bouter, L. M., Veldkamp, C., Van de Ven, P. M., Wicherts, J. M., & Smulders, Y. M. (2016). Personality traits are associated with research misbehavior in Dutch scientists: a cross-sectional study. *PLoS One*, 11(9), e0163251.
- Toomela, A. (2019). *The psychology of scientific inquiry*. Cham: Springer.
- Toth-Kiraly, I., Bothe, B., & Orosz, G. (2017). Exploratory structural equation modeling analysis of the Self-Compassion Scale. *Mindfulness*, 8, 881–892.
- Ursic, N., Kocjancic, D., & Zvelc, G. (2019). Psychometric properties of the Slovenian long and short version of the Self-Compassion Scale. *Psihologija*, 52, 107–125.
- Van Dam, N. T., Sheppard, S. C., Forsyth, J. P., & Earleywine, M. (2011). Self-compassion is a better predictor than mindfulness of symptom severity and quality of life in mixed anxiety and depression. *Journal of Anxiety Disorders*, 25, 123–130.
- Wadsworth, L. P., Forgeard, M., Hsu, K. J., Kertz, S., Treadway, M., & Bjorgvinsson, T. (2018). Examining the role of repetitive negative thinking in relations between positive and negative aspects of self-

- compassion and symptom improvement during intensive treatment. *Cognitive Therapy and Research*, 42, 236–249.
- 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.
- Wilson, A. C., Mackintosh, K., Power, K., & Chan, S. W. Y. (2019). Effectiveness of self-compassion related therapies: a systematic review and meta-analysis. *Mindfulness*, 10, 979–995.
- Wong, N., Rindfleisch, A., & Burroughs, J. E. (2003). Do reverse worded items confound measures in cross-cultural consumer research? The case of the Material Values Scale. *Journal of Consumer Research*, 30, 72–91.
- Yarnell, L. M., Stafford, R. E., Neff, K. D., Reilly, E. D., Knox, M. C., & Mullarkey, M. (2015). Meta-analysis of gender differences in self-compassion. *Self and Identity*, 14, 499–520.
- Yarnell, L. M., Neff, K. D., Davidson, O. A., & Mullarkey, M. (2019). Gender differences in self-compassion: examining the role of gender role orientation. *Mindfulness*, 10, 1136–1152.
- Zessin, U., Dickhäuser, O., & Garbade, S. (2015). The relationship between self-compassion and well-being: a meta-analysis. *Applied Psychology: Health and Well-being*, 7, 340–364.
- Zhang, H., Dong, L., Watson-Singleton, N. N., Tarantino, N., Carr, E. R., Niles-Carnes, L. V., Patterson, B., & Kaslow, N. J. (2019). Psychometric properties of the Self-Compassion Scale in an African American sample. *Mindfulness*, 10, 1395–1405.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.