CReaTE

Canterbury Research and Theses Environment

Canterbury Christ Church University's repository of research outputs

http://create.canterbury.ac.uk

Please cite this publication as follows:

Link to official URL (if available):

Contact: create.library@canterbury.ac.uk

Canterbury
Christ Church
University

Accepted Manuscript

What is Compassion and How Can We Measure it? A Review of Definitions and Measures

Clara Strauss, Billie Lever Taylor, Jenny Gu, Willem Kuyken, Ruth Baer, Fergal Jones, Kate Cavanagh

What is Compassion and How Can We Measure it? A Review of Definitions and Measures

Clara Strauss[1,2,*]; Billie Lever Taylor[2,3]; Jenny Gu[1], Willem Kuyken[4], Ruth Baer[4,5],

Fergal Jones[6] & Kate Cavanagh[1]

[1] School of Psychology, University of Sussex, Falmer, East Sussex, BN1 9QH, UK

[2] Sussex Partnership NHS Foundation Trust, Hove, BN3 7HZ, UK

[3] Division of Psychiatry, University College London, 149 Tottenham Court Road, London, W1T 7NF, UK

[4] Department of Psychiatry, University of Oxford, Warneford Hospital, Headington, Oxford, OX3 7JX, UK

[5] Department of Psychology, University of Kentucky, Lexington, KY 40506-0044, US

[6] Salomons Centre for Applied Psychology, Canterbury Christ Church University, Runcie Court, Broomhill Road, Tunbridge Wells, TN3 0TF, UK

[*]Correspondence concerning this article should be addressed to:

Clara Strauss, School of Psychology, University of Sussex, Falmer, Brighton, BN1 9QJ, UK

Email: c.y.strauss@sussex.ac.uk

Phone: 01273 265896

Abstract

The importance of compassion is widely recognized and it is receiving increasing research attention. Yet, there is lack of consensus on definition and a paucity of psychometrically robust measures of this construct. Without an agreed definition and adequate measures, we cannot study compassion, measure compassion or evaluate whether interventions designed to enhance compassion are effective. In response, this paper proposes a definition of compassion and offers a systematic review of self- and observer-rated measures. Following consolidation of existing definitions, we propose that compassion consists of five elements: recognizing suffering, understanding the universality of human suffering, feeling for the person suffering, tolerating uncomfortable feelings, and motivation to act/acting to alleviate suffering. Three databases were searched (Web of Science, PsycInfo, and Medline) and nine measures included and rated for quality. Quality ratings ranged from 2 to 7 out of 14 with low ratings due to poor internal consistency for subscales, insufficient evidence for factor structure and/or failure to examine floor/ceiling effects, test-retest reliability, and discriminant validity. We call for empirical testing of our five-element definition, if supported, the development of a measure of compassion based on this operational definition, and which demonstrates adequate psychometric properties.

Keywords: compassion; self-compassion; measure; systematic review; definition

Introduction

The importance of compassion is recognized in many segments of society. Most of the world's religious traditions place compassion at the center of their belief systems. International professional bodies in healthcare, education and the justice system also emphasize the importance of compassion. In the US, compassion is enshrined in the American Medical Association's (AMA) Principles of Medical Ethics, with Item 1 stating that "A physician shall be dedicated to providing competent medical services with compassion and respect for human dignity" (AMA, 1981). In the UK, compassion is one of the six core values in the NHS constitution (Department of Health; DoH, 2013), and calls for a greater focus on compassion have been driven in part by high profile exposés of serious failings in compassionate care at some hospitals and care homes. The international 'Compassion in Education' foundation (CoED, 2014) offers a range of services to educational professionals in order to promote compassion in the education system. It has also been argued that compassion should lie at the core of the ethical framework guiding our justice systems (Norko, 2005).

An evolutionary perspective on compassion can be traced to Darwin (1871), who stated that "those communities which included the greatest number of the most sympathetic members would flourish best, and rear the greatest number of offspring" (p. 130). Current theorists also note that compassion is reproductively advantageous, being part of the care-giving system that has evolved to nurture and protect the young (e.g. Gilbert, 2005; Goetz, Keltner, & Simon-Thomas, 2010). Compassion can be seen as having evolved from an adaptive focus on protecting oneself and one's offspring to a broader focus on protecting others including and beyond one's immediate kinship group (de Waal, 2009). Compassion may also have evolved in primates because it is a desirable criterion in mate selection and facilitates cooperative relationships with non-kin (e.g. de Waal, 2009; Keltner, 2009).

3

Within the healthcare domain, compassion is believed to have numerous practical advantages. It has been argued that treating patients compassionately has wide-ranging benefits, including improving clinical outcomes, increasing patient satisfaction with services, and enhancing the quality of information gathered from patients (Epstein et al., 2005; Rendelmeir et al., 1995; Sanghavi, 2006). Conversely, compassion fatigue may contribute to poor quality of care (Najjar et al., 2009). Treating oneself and others with compassion is also believed to promote individual wellbeing and improve mental health (e.g. Cosley et al., 2010; Feldman & Kuyken, 2011; MacBeth & Gumley, 2012). Accordingly, some researchers have called for the implementation of interventions that seek to enhance people's ability to give and receive compassion (e.g. Gilbert, 2005; 2010; Neff & Germer, 2012), arguing that compassion buffers reactivity to stress and is central to the process of recovery from psychopathology. Other research has focused on the developmental trajectory of compassion and has found relationships between parenting styles and children's levels of sympathy and caring (Eisenberg et al., 2015) and between attachment security in childhood and capacity for compassion in adulthood (see Gillath, Shaver, & Mukilincer, 2005, for a review).

Despite the importance of compassion and increasing interest from researchers, clinicians, teachers, and other professionals, there is lack of consensus on its definition and a paucity of psychometrically robust measurement tools. Without these, scientific enquiry is greatly impeded – we need consensus on a definition and valid and reliable measurement tools in order to assess compassion in empirical research. This paper has two aims: first, to suggest a definition of compassion based on a consolidation of conceptualizations and definitions in the field and second, to systematically review self- and observer-rated measures of compassion.

Conceptualizations of Compassion: Towards a Definition

4

According to the Oxford English Dictionary, the word "compassion" stems from the Latin "compati", meaning "to suffer with". In the literature, there appears to be a broad consensus that compassion involves feeling for a person who is suffering and being motivated to act to help them (e.g. Lazarus, 1991; Goetz et al., 2010). For example, in his seminal work on human emotions Lazarus defines compassion as: "Being moved by another's suffering and wanting to help" (p. 289). Similarly, in a major systematic review of compassion and its evolutionary origins, Goetz et al. define it as: "the feeling that arises in witnessing another's suffering and that motivates a subsequent desire to help" (p. 351). These definitions have in common the suggestion that compassion is not only about feeling touched by a person's suffering, but also about wanting to act to help them. Compassion is a fundamental tenet of Buddhist philosophy (it is, in fact, emphasized by all the main world religions but Buddhist perspectives on compassion have been given greater prominence in the psychological literature) and the Dalai Lama (1995) defines compassion in comparable terms as: "An openness to the suffering of others with a commitment to relieve it". However, within Buddhism, compassion is seen not only as an emotional response but also as a response founded on reason and wisdom which is embedded in an ethical framework concerned with the selfless intention of freeing others from suffering.

More specifically, in their review of compassion within organizations, Kanov et al. (2004) argue that compassion consists of three facets: noticing, feeling, and responding. 'Noticing' involves being aware of a person's suffering, either by cognitively recognizing this suffering or by experiencing an unconscious physical or affective reaction to it. 'Feeling' is defined as responding emotionally to that suffering and experiencing 'empathic concern' through adopting the person's perspective and imagining or feeling their condition. Finally, 'responding' involves having a desire to act to alleviate the person's suffering. As in Buddhist conceptualizations, this definition suggests that compassion does not purely consist

of affective and behavioral elements, but also may have cognitive components insofar as it involves being able to imagine and reason about a person's experiences.

Gilbert (2010) conceptualizes compassion in evolutionary terms, arguing that compassion is an evolved motivational system designed to regulate negative affect, where compassion is seen to have originated from the same capacities that primates evolved to form attachment bonds and engage in affiliative and cooperative behaviors for group survival. He defines compassion as: "A deep awareness of the suffering of another coupled with the wish to relieve it" (Gilbert, 2009, p. 13) and, like Kanov et al. (2004), suggests it has cognitive, affective and behavioral elements. Gilbert (2010) sees compassion as consisting of six 'attributes': sensitivity, sympathy, empathy, motivation/caring, distress tolerance and non-judgement. 'Sensitivity' involves being responsive to other people's emotions and perceiving when they need help, which appears to correspond to Kanov et al.'s 'noticing' facet. 'Sympathy' (defined as showing concern for the other person's suffering) and 'empathy' (defined as putting yourself in their shoes) together appear to correspond to Kanov et al.'s 'feeling' facet. Finally, 'motivation' to act is akin to Kanov et al.'s 'responding' facet.

The final two components in Gilbert's (2010) model – 'distress tolerance' and 'non-judgement' – are not included in Kanov et al.'s (2004) model. Distress tolerance is defined as the ability to tolerate difficult emotions in oneself when confronted with someone else's suffering without becoming overwhelmed by them. Gilbert argues that this is important because if we over-identify with a person's suffering we may feel a need to get away from them or reduce our awareness of their distress, preventing a compassionate response. This suggests that, although compassion is about 'suffering with' another person, if we feel such extreme personal distress in the face of another's suffering that we become too focused on our own discomfort, this may hinder our ability to help. The final element of Gilbert's model – 'non-judgement' – is defined as the ability to remain accepting of and tolerant towards

6

another person even when their condition, or response to it, gives rise to difficult feelings in oneself, such as frustration, anger, fear or disgust. The idea that compassion means approaching those who are suffering with non-judgement and tolerance – even if they are in some sense disagreeable to us – is also central to Buddhist conceptualizations. For example, the Dalai Lama (2002) contends that: "for a practitioner of love and compassion, an enemy is one of the most important teachers. Without an enemy you cannot practice tolerance, and without tolerance you cannot build a sound basis of compassion" (p. 75).

Both Gilbert (2005, 2010) and the Dalai Lama are also clear that compassion is not only felt for close others (where attachment comes into play as well), but also for those we do not know. Similarly, Gilbert (2003, cited in Wang, 2005) notes: "One can feel compassion for those we might never meet (the starving children in Africa)" (p. 99-100). The idea that compassion can be experienced towards close others and those we do not know is also emphasized by Sprecher and Fehr (2005) who developed a measure of 'compassionate love' which includes separate versions relating to close others and strangers or humankind at large.

Like Gilbert (2010), Wispe (1991) conceptualizes compassion for others not only as being aware of and moved by suffering and wanting to help, but also as including the ability to adopt a non-judgmental stance towards others and to tolerate one's own distress when faced with other people's suffering. Neff (2003a) developed this definition of compassion for others into a model of self-compassion, arguing that self-compassion can be viewed as compassion directed inward towards the self. She concludes that self-compassion consists of three principal components: kindness (being kind and non-judgmental towards the self rather than self-critical), mindfulness (which, like 'distress tolerance', involves holding painful feelings in mindful awareness rather than over-identifying with them), and common humanity (seeing one's suffering as part of the human condition rather than as isolating).

It is debatable whether compassion for others and self-compassion are in fact part of the same overarching construct. While Buddhist thinking argues that differentiating compassion for others from self-compassion means drawing a false distinction between the self and others, and moreover that self-compassion is a prerequisite for showing 'true' compassion towards others, recent research has found that associations between self-compassion and other-focused compassion may be weak, or even non-existent for some populations. For example, Neff and Pommier (2013) explored the relationship between self-compassion and compassion for others and found that they were not correlated in a sample of undergraduates (r = .00), and only weakly correlated in a community sample and a sample of practicing meditators (r = .15 and .28 respectively). Similarly, Pommier (2010) found no association between self-compassion and compassion for others in a sample of undergraduates (r = .07). It is unclear whether the lack of association between self-compassion and compassion for others reflects a genuine independence between these two constructs or whether it reflects definitional problems, weaknesses of correlational study designs or limitations with current measures (e.g. Williams, Dalgleish, Karl, & Kuyken, 2014). This is an area for further empirical research.

While acknowledging some of the difficulties with equating self-compassion with compassion for others, Pommier (2010) has applied Neff's (2003a) model of self-compassion to a model of compassion for others suggesting that, like self-compassion, compassion for others can be seen as involving kindness, mindfulness and common humanity. In Pommier's model, 'kindness' is defined as being understanding towards others who are suffering instead of being critical or indifferent towards them. 'Mindfulness' is seen as the ability to notice another person's suffering and remain open to it without feeling so distressed that you disengage from that person. And 'common humanity' is conceptualized as realizing that all

8

humans suffer and that one could find oneself in the position of the sufferer if one was less fortunate – a sense that "There but for the grace of God, go I".

This emphasis on seeing a 'common humanity' with the person who is suffering is also evident in Buddhist definitions of compassion, with the Dalai Lama (2005) arguing that: "Genuine compassion must have both wisdom and loving kindness. That is to say, one must understand the nature of the suffering from which we wish to free others (this is wisdom), and one must experience deep intimacy and empathy with other sentient beings (this is loving kindness)" (p. 49). Within such Buddhist conceptualizations, understanding the nature of suffering ('wisdom') is to understand that suffering is part of what it is to be human; suffering is a shared human experience. Similarly, in their review of the role of compassion in mindfulness-based therapies, Feldman and Kuyken (2011) describe compassion as: "an orientation of mind that recognizes pain and the universality of pain in human experience and the capacity to meet that pain with kindness, empathy, equanimity and patience" (p. 145).

In summary, in all these definitions compassion is seen as awareness of someone's suffering, being moved by it (emotionally and, according to some definitions, cognitively), and acting or feeling motivated to help. Several definitions emphasize that, although one is moved by suffering, compassion also involves being able to tolerate uncomfortable feelings that arise in oneself as a result of seeing suffering, including tolerating feelings of distaste, frustration or anger that might be elicited by that suffering. There is also a suggestion in several definitions that compassion involves recognizing a commonality with the sufferer, acknowledging that as a fellow being we too could find ourselves in a similar position. Table 1 contains summaries of the major definitions of compassion discussed in this section.

Related Constructs

9

In definitions of compassion, reference is commonly made to related terms such as empathy and in turn these words are often used to define each other. The similarities between compassion and constructs such as kindness, pity and altruism have also been noted (Goetz et al., 2010). It is instructive to consider the overlaps and distinctions between these terms.

According to the Oxford English Dictionary, the word empathy is defined as: "the power of mentally identifying oneself with (and so fully comprehending) a person or object of contemplation". Like compassion, empathy has been described as a multidimensional construct, consisting of cognitive and affective components (Davis, 1983). Cognitive empathy can be defined as intellectually understanding another person's emotions and perspective (Hogan, 1969), whereas affective empathy refers to being affected by and sharing another's emotions (Mehrabian & Epstein, 1972). Gilbert (2010), Kanov et al. (2004) and the Dalai Lama (2005) all explicitly define compassion as requiring empathy and therefore appear to see empathy as an essential element of compassion. Even so, they suggest that compassion has additional components over and above empathy. In particular, a desire to act or acting to alleviate suffering is seen as a core feature of compassion but not empathy (see Table 1).

A second distinction between compassion and empathy is that, whereas compassion is felt specifically in response to suffering, empathy may apply to a broader range of situations, for example one could feel empathy with someone else's anger, fear or joy (Pommier, 2010). Moreover, Goetz et al. (2010) argue that compassion is an emotion in its own right, whereas empathy is the vicarious experience of another's emotions, while Sprecher and Fehr (2005) contend that compassion is broader than empathy because it can be felt for humanity at large, rather than only in relation to specific interpersonal encounters. In addition, recent neuroscientific findings suggest that different brain regions are activated in response to compassion and empathy training (Klimecki, Leiberg, Ricard, & Singer, 2014).

The same is true of pity, which, despite also having similarities to compassion, does not require an inclination to help. On the contrary, some writers have argued that pity implies that one sees someone as unworthy of help (Lazarus, 1991), or at least involves showing condescension towards them (e.g. Cassell, 2002). At the other end of the spectrum, altruism has a greater focus than compassion on behavioral acts that may be at a great personal cost to the person. Also, altruistic acts can have a broad range of motivations, that do not necessarily involve the same elements as compassion. .

Finally, compassion is frequently linked to kindness (defined by the Oxford English Dictionary as "the quality of being friendly, generous and considerate"). For example, Neff (2003a) and Pommier (2010) include 'kindness' as a component of compassion and compassion has even been defined as "intelligent kindness" (e.g. DoH, 2013). However, these two terms have distinctions too: for example, as outlined compassion includes elements beyond kindness (e.g. recognizing and being touched by suffering); and likewise, kindness includes elements beyond compassion, as kindness is not only linked to suffering (e.g. remembering someone's birthday is kind but not compassionate). Additionally, compassion may not always involve kindness in the moment (e.g. taking a 'tough love' approach may be compassionate but not kind). .

Compassion: A Proposed Definition

To bring together the various definitions and considerations above and to aid the review of existing measures of compassion, we propose a new definition of compassion as a cognitive, affective, and behavioral process consisting of the following five elements that refer to both self- and other-compassion: 1) Recognizing suffering; 2) Understanding the universality of suffering in human experience; 3) Feeling empathy for the person suffering and connecting with the distress (emotional resonance); 4) Tolerating uncomfortable feelings

11

aroused in response to the suffering person (e.g. distress, anger, fear) so remaining open to and accepting of the person suffering; and 5) Motivation to act/acting to alleviate suffering.

We use this proposed new definition of compassion to organize the remainder of this paper, which provides a systematic review of self- and observer-rated measures of compassion. The psychometric properties of identified measures are reported and rated for quality, including the extent to which they measure each of the five elements outlined above.

Method

Inclusion and Exclusion Criteria

To be included in the main review, measures had to: be available in English; include a scale explicitly defined by its authors as measuring compassion; include a psychometric paper outlining the development of the scale; and be obtainable either within a published article or from the author (two attempts were made to contact the relevant authors to obtain measures where necessary). Measures were excluded if they did not assess participants' levels of compassion per se (e.g. measures of barriers to feeling compassion, fear of compassion, and empathy were excluded); used non-questionnaire measures of compassion, or included only a subscale on compassion. Because we do not yet know the relationship between compassion for others and self-compassion, measures of self-compassion were included because many conceptualizations and definitions of compassion do not distinguish between other- and self-compassion.

Information Sources

The databases searched for relevant measures included Web of Science (Thomson Reuters), PsycInfo, and Medline, from inception to 23 September 2015. Dissertations and

theses that met the inclusion criteria were reviewed along with papers published in peer-reviewed journals. Where relevant, the most recent versions of measures were reviewed.

Search Strategy

All articles including the word "compassion*" in combination with "measure*", "scale*", "instrument*" or "questionnaire*" in either the title or abstract or key words were identified. Where identified papers referred to additional scales, reference lists were searched and any additional relevant papers retrieved. Experts in the field were also consulted to ensure that no measures were missed.

Assessment of Quality

The psychometric properties of each measure were reviewed and measures were rated for quality based largely on Terwee et al.'s (2007) quality criteria for health status measures. These criteria were used because they include explicit criteria for what constitutes good measurement properties. However, since these criteria relate to measures of health status, Barker, Pistrang, and Elliott's (2002) 'rules of thumb' for evaluating psychological measures were also drawn on where these seemed more appropriate[1]. Terwee et al. award measures a positive (+), intermediate (?), or negative (-) rating, or a rating of 0 where no information regarding the relevant criteria is provided. In this review, in order to make scores easier to interpret, measures were given a score of 2 if there was evidence for the criterion being fully met, 1 if the criterion was partially met, and 0 if the criterion was not met, or if no relevant data were reported. Scores were aggregated to provide an overall rating. Two researchers

---

[1] Terwee et al. (2007) proposed the following eight quality criteria to evaluate health status measures: 1) Content validity, 2) internal consistency, 3) criterion validity, 4) construct validity (convergent and discriminant validity), 5) reproducibility (test-retest reliability), 6) responsiveness, 7) floor and ceiling effects, and 8) interpretability. We did not include criterion validity and responsiveness as criteria, for the reasons stated in the paper. Terwee et al. did not provide rules of thumb in terms of the size of correlation coefficients for the test-retest reliability criterion. They also did not account for the size of correlations for the convergent and discriminant validity criterion. Therefore, for these two criteria, we drew on Barker at al.'s (2002) general recommendations when evaluating the reliability and validity of psychological measures. We also included factor structure as a criterion.

independently scored the measures using these criteria, and any discrepancies in scoring were resolved collectively. Specifically, measures were rated across the following domains:

1. Content validity (the extent to which the domain of interest was comprehensively sampled by the items in the questionnaire). In this case the domain of interest was considered to be compassion as defined in this review, rather than as defined by the scale's authors. Under this criterion, Terwee et al. (2007) also emphasize the importance of both members of the target population and experts being involved in item development. For this criterion to be fully met all five elements of compassion must be captured by the items and items must have been generated in consultation with experts and members of the intended population.

   2. Factor structure (whether or not the factor structure for the measure has been examined and supported). This criterion was included in addition to those proposed by Terwee et al. (2007). This criterion was scored as follows. A score of 2 was given where exploratory factor analysis (EFA) followed by confirmatory factor analysis (CFA) have been conducted in independent samples OR where CFA has been conducted if the factor structure has been previously proposed theoretically (a score of 2 was only given if the factor analyses support the proposed factor structure). A score of 1 was given if only EFA has been conducted (without CFA) and if the EFA supports the factor structure. A score of 0 was given where either factor analysis has not been conducted OR where EFA and/or CFA have been conducted that do not support the proposed factor structure.

3. Internal consistency (the extent to which items in a (sub)scale are inter-correlated and thus measuring the same construct). For this criterion to be fully met - in line with Terwee et al.'s (2007) criteria - factor analyses had to have been performed on an adequate sample

14

size (7 * number of items and N > 100) and Cronbach's alpha for each identified factor

had to be between .70 and .95.

4. Test-retest reliability. Based on Barker et al.'s (2002) 'rules of thumb' test-retest

   reliabilities had to be at least r = .70 for this criterion to be fully met.

5. Convergent and discriminant validity (the extent to which scores on a particular scale

   relate to other measures in a manner consistent with theoretically derived hypotheses). For

   this criterion to be met, Terwee et al. (2007) require that (i) specific hypotheses are

   formulated by the scale's authors about expected correlations and (ii) at least three

   quarters of results are in line with expectations. As Terwee et al. do not take into account

   the strength of these correlations, we also drew on Barker et al. (2002), and required that at

   least two correlations with theoretically related constructs were at least r = .50 to

   demonstrate convergent validity.

6. Floor and ceiling effects (the number of respondents achieving the highest or lowest

   possible scores). This was rated based on Terwee et al.'s (2007) criterion that no more

   than 15% of the sample should receive the top or bottom score on a scale.

7. Interpretability (how differences in scores on the measure can be interpreted, or the degree

   to which qualitative meaning can be attached to quantitative scores). Terwee et al. (2007)

   require means and SDs of scores from at least four relevant subgroups of participants to be

   reported (e.g. compassion scores in males vs. females, meditators vs. non-meditators) and

   minimal important change defined. However, as minimal important change was arguably

   not entirely relevant to the measures in this review, consideration was instead given to

   whether the authors indicated how scale scores might be interpreted.

Terwee et al.'s (2007) quality ratings also include 'criterion validity' (the extent to which scores on a particular scale relate to a 'gold standard') and 'responsiveness' (the ability of a scale to detect change over time). However, these two criteria were not rated. In the case of 'criterion validity' this was because there is no gold standard compassion measure to rate scales against. In the case of 'responsiveness' it was because such data were not typically available and, as this criterion relates to clinically meaningful change, arguably the majority of the scales were not primarily designed to measure this. Therefore, the total possible score for any measure was 14.

Results

Review of Identified Measures

Figure 1 shows a flow diagram illustrating the search process. After removing duplicates, 2,146 papers were identified, with only nine measures included after screening titles, abstracts, and full texts. Table 2 provides the quality ratings of the reviewed measures and Tables 3 and 4 outline the psychometric properties of the measures. Floor and ceiling effects are not included in Tables 3 and 4 because no studies reported them. Similarly, although most studies included measures of related constructs to test convergent validity, none included measures of theoretically unrelated constructs; therefore, discriminant validity is not included in Table 4.

Compassionate love scale (CLS; Sprecher & Fehr, 2005). The CLS consists of 21 self-report items, rated on a Likert scale from 1 (not at all true of me) to 7 (very true of me). The CLS is intended for the general population and consists of two forms: one relating to significant others (including family members and friends) and one focusing on strangers and humanity at large.

Content validity. The scale was rated as partially satisfactory for content validity. Items were generated by the investigators based on a review of the literature on love and altruism and also based on a prototype analysis with laypeople around their concept of compassionate love. In line with our definition of compassion, the scale includes items related to four of our five elements of compassion identified earlier: Feeling moved by other people's suffering (emotional resonance), understanding or imagining something about their condition as a fellow being, accepting and not judging others (which implies tolerance), and being motivated to help them. However, the CLS did not appear to contain items explicitly related to recognizing suffering.

Three items include the word 'compassion' or 'compassionate love', which requires respondents to define these concepts themselves. However, it seems uncertain whether they will know what is meant by 'compassionate love' or define it uniformly. Additionally, not all items on the scale relate to those who are suffering, and it is questionable whether items such as: "I feel happy when I see that [loved ones/others (strangers)] are happy" and "I very much wish to be kind and good to [my friends and family members/fellow human beings]" assess compassion or in fact more broadly assess empathy and kindness respectively. Finally, the scale refers explicitly to either close others or strangers, but does not allow respondents to consider people who may not fall into either of these categories (e.g. patients responding in relation to healthcare professionals), potentially limiting its use in some contexts.

Factor structure and reliability. Exploratory factor analysis (EFA) yielded a single factor structure for each version of the scale. Sprecher and Fehr (2005) did not explicitly propose a factor structure for the CLS prior to analysis and did not conduct CFA. Internal consistency was high for both versions. Test-retest reliability was not reported.

17

Convergent validity and interpretability. Convergent validity was supported by significant correlations in the expected directions with measures of empathy, helpfulness, volunteerism, and spiritual experiences. Limited subgroup analyses were undertaken by Sprecher and Fehr (2005), showing that women obtained significantly higher compassion scores than men on both versions.

Santa Clara brief compassion scale (SCBCS; Hwang, Plante, & Lackey, 2008). The SCBCS is a shortened version of Sprecher and Fehr's (2005) CLS, consisting of five items from the original scale (the correlation between the two scales is r = .95). Unlike the CLS, this scale refers exclusively to strangers rather than to close others. The items of the SCBCS were selected because they had moderate means, high standard deviations, and high correlations with the overall score from the CLS.

Content validity. The SCBCS was rated partially satisfactory for content validity. The scale includes items related to two of our five elements of compassion: Emotionally connecting with other people's suffering and acting to help them. However, unlike the CLS, the SCBCS did not appear to contain items explicitly related to understanding the universality of suffering and tolerating uncomfortable feelings, and also did not include items explicitly related to recognizing suffering. Two items contain the word "compassion", again relying on respondents to define these terms rather than tapping into their underlying elements.

Factor structure and reliability. EFA yielded a single factor structure for the SCBCS and CFA was not conducted. Internal consistency was high. Test-retest reliability was not reported.

Convergent validity and interpretability. The SCBCS was strongly correlated with empathic concern, moderately correlated with vocational identity, and showed a small

correlation with strength of religious faith. Examination of group differences was limited to gender and showed that women scored significantly higher than men.

The compassion scale (CS-M; Martins, Nicholas, Shaheen, Jones, & Norris, 2013). Martins et al.'s CS-M is a 10-item self-report scale developed to measure five domains of compassion: generosity, hospitality, objectivity, sensitivity, and tolerance across social networks and relationships (strangers, friends, and family) using a 1 (none) to 7 (all) response scale. The aim of the scale was to provide a measure of compassion across domains that could be enhanced through training, as the authors argue that scales like the CLS do not lend themselves well to measuring compassion in a way that can be targeted for education. Items were generated and evaluated by a panel of academic and community experts.

Content validity. Martins et al.'s (2013) scale was rated partially satisfactory for content validity. The CS-M focuses exclusively on practical acts of compassion including giving financial help to others, using your free time to help others, and doing things for others at a cost or risk to yourself or your family and friends. Thus, only the acting to alleviate suffering factor of our five-factor definition is captured by the items of the CS-M; items related to recognizing suffering, understanding the universality of suffering, emotional resonance, and tolerating uncomfortable feelings were not included. Additionally, it could be argued that the scale's items measure only a limited range of acts of compassion (giving away money, using free time to help others, sharing personal space, or doing something for others at a cost to oneself) and if the scale were applied to certain contexts (e.g. a healthcare context), the items may not assess the types of actions that might be expected in those contexts (e.g. considering ways to make those who are suffering more comfortable). Indeed, it is not altogether clear for what population the scale is intended. Furthermore, items such as "How many times would you do the right thing if it puts your family at risk" do not appear to fit well with a response ranging from "none" to "all".

Factor structure and reliability. EFA did not support the proposed five factor structure; the analysis identified a two-factor solution. However, the two-factor structure was rejected by the authors in favour of a single factor model, arguing that, as all items beginning "How much of your…?" loaded onto one factor and all items beginning "How many times would you…?" loaded onto the second factor, the factors appeared to reflect methodological differences between items rather than substantively different constructs. Cronbach's alpha for the total scale was acceptable. Test-retest reliability was not tested.

Convergent validity and interpretability. The authors only compared their scale with the CLS (r > .50). In terms of interpretability, the authors provided mean scores for a range of subgroups but, though they argue that the scale should help measure change in compassion after training, they do not provide any indication of what level of change on the scale would be needed to show that such training had been of value.

Self-compassion scale (SCS; Neff, 2003b). The SCS is a 26-item scale with a 5-point response scale from "almost never" to "almost always".

Content validity. The scale was rated partially satisfactory overall for content validity. Although items were selected after extensive piloting, it is notable that this was only carried out with experts and undergraduate students, even though the scale's target population included community and clinical samples. The scale includes items related to four of the five elements in the definition of compassion used in this review: understanding the universality of suffering, emotional resonance, the ability to tolerate distressing feelings, and feeling motivated to act or acting to help ameliorate one's suffering. However, the scale does not include items specifically relating to being attentive to how one is feeling.

Factor structure and reliability. CFA of the SCS supported the six factor model, with each of the three components of self-compassion split into two sub-factors - one comprising

'positively' worded and one 'negatively' worded items. This resulted in the following factors: kindness versus self-judgement; mindfulness versus over-identification; and common humanity versus isolation. However, Neff (2003b) found only a marginal fit with a single higher-order factor, questioning whether the six factors can be explained by a single overarching construct of self-compassion. Other studies have also questioned the higher-order factor structure and the non-hierarchical six-factor model across a range of populations, student, clinical and meditating/non-meditating (e.g. Costa, Marôco, Pinto-Gouveia, Ferreira, & Castilho, 2015; López et al., 2015; Williams et al., 2014). Several studies have suggested a two-factor model of self-compassion, with the factors representing the positive and negative dimensions of self-compassion and self-criticism, respectively (e.g. Costa et al., 2015; López et al., 2015).

Recently, Neff (2015) argued that the two-factor conceptualization of the SCS is problematic in that it does not capture the relative balance between the three proposed broad components of self-compassion (self-kindness vs. self-judgment, common humanity vs. isolation, and mindfulness vs. over-identification). Instead, Neff proposed a bifactor model of self-compassion, where each item loads directly on to a general factor as well as their respective subscale, and suggests that researchers can select whether to analyse subscale scores separately or use a total SCS score depending on their interests.

Cronbach's alpha values for total SCS scale and subscale scores and test-retest reliability were adequate.

Convergent validity and interpretability. Convergent validity was supported by significant correlations in the expected direction between the SCS and other related measures, several of which were ≥ .50. A partially satisfactory score was achieved for interpretability, as only gender differences were reported.

Self-compassion scale: Short form (SCS-SF; Raes, Pommier, Neff, & Van Gucht, 2011). Raes et al. developed a 12-item version of the SCS by selecting two items from each of the SCS's six subscales, based on their high correlations with the SCS and intended subscales, and high intercorrelations. The SCS-SF is rated in the same way as the SCS.

Content validity. The scale was rated 'partially satisfactory' for content validity for the same reasons as the long form.

Factor structure and reliability. CFA supported the proposed six-factor hierarchical structure of the measure. Internal consistency was acceptable for the total score, but was variable for the individual subscales. Test-retest reliability was not reported.

Convergent validity and interpretability. Relevant data were not reported for convergent validity and interpretability.

The compassion scale (CS-P; Pommier, 2010). The CS-P is a 24-item self-report scale targeted at the general population and based on the argument (outlined earlier) that compassion consists of six elements: Kindness (in contrast to indifference), mindfulness (in contrast to disengagement), and common humanity (in contrast to separation). Responses are given on a five-point Likert scale, ranging from 1 (almost never) to 5 (almost always).

Content validity. The CS-P was rated partially satisfactory for content validity. Items were devised by the author, based on theory and research, and reviewed by a panel of experts. The scale includes items consistent with four of our five elements of compassion: Recognizing suffering, feeling moved by suffering, understanding or imagining something about another person's condition as a fellow being, and motivation to act/acting to alleviate suffering. Although in her development of the scale Pommier (2010) notes that compassion requires the ability to tolerate uncomfortable feelings in the face of suffering so that one can

22

remain tolerant and accepting of others, the scale appears not to directly assess this, other than asking whether respondents "try to keep a balanced perspective on the situation" when people tell them about their problems, or whether they tend to avoid those who suffer.

Additionally, several of the scale's items include the words 'sometimes', 'often' or 'usually' which conflict with the response scale used ('almost never' to 'almost always') and makes responses difficult to interpret. The response scale is perhaps also unintuitive for negatively worded items – for example, a response of "I almost never don't feel emotionally connected to people in pain" may be difficult for some people to rate accurately. Similarly, items such as "Suffering is just a part of the common human experience" cannot be answered accurately using the scale from 'almost never' to 'almost always' and do not sit well with the scale's instructions to "indicate how often you behave in the stated manner".

Factor structure and reliability. CFA supported the proposed six-factor structure of the measure, and that a single higher order factor of compassion explained the inter-correlations between the six factors. EFA was not conducted because Neff's (2003b) Self-Compassion Scale (SCS) had already demonstrated these six factors. However, as noted earlier compassion for others and self-compassion were not significantly correlated in Pommier's (2010) research, suggesting that the factor structure for each measure cannot be assumed to be identical. Internal consistency was high for the total score but mixed and inconsistent across samples for the subscales. Test-retest reliability was not reported.

Convergent validity and interpretability. Convergent validity was supported by significant correlations in the expected direction between the CS-P and other measures of compassion, empathy, perspective-taking, and wisdom; several of these were $\geq .50$. However, the CS-P was not significantly correlated with the SCS (Neff, 2003b), a problematic finding for the scale's construct validity, given that the CS-P was developed based on the factor

23

structure of the SCS. Additionally, while the scale was positively correlated with the CLS, this correlation was only small for the strangers-humanity version (r = .27; r = .54 for the close others version). This suggests that the CS-P and the CLS may not be measuring the same construct. Another unexpected finding was that the Southampton Mindfulness Questionnaire (SMQ; Chadwick et al., 2008) had a small negative correlation with the CS-P. The only subgroup analyzed was gender, again showing that women scored higher than men.

Relational compassion scale (RCS; Hacker, 2008). The RCS consists of 16 items rated on a four-point scale (from 'do not agree' to 'agree strongly'). The scale consists of four subscales which measure respondents' compassion for others and self-compassion, along with their beliefs about how compassionate other people are to each other, and their beliefs about how compassionate other people are to them. The latter two subscales extend beyond simply measuring respondents' own levels of compassion, but the scale was nevertheless included because it defines itself as a comprehensive measure of compassion and also because the subscales were psychometrically tested individually.

Content validity. The RCS was rated 'partially satisfactory' for content validity. The scale's items that comprise the 'compassion for others' subscale assess people's capacity to recognize and understand suffering and accept and not judge others (which implies tolerance), just two of the five elements in our definition of compassion. Additionally, some items relate to other people's 'experiences' in general, rather than specifically to their suffering. The items comprising the self-compassion subscale assess emotional resonance and acting to alleviate suffering, two of the five elements in our definition. Items related to understanding the universality of suffering were not included in either subscale.

Factor structure and reliability. CFA using the final version of the RCS supported its proposed four-factor structure. Internal consistency was acceptable for all four subscales. Test-retest reliability was not tested.

Convergent validity and interpretability. Although several correlations with related measures were ≥ .50, specific hypotheses appear not to have been set out in advance about the expected direction of correlations and, in the discussion, the author highlights some unexpected findings. For example, the 'compassion for others' subscale did not correlate significantly with a measure of self-criticism/self-attack and self-reassurance (Gilbert, Clarke, Hempel, Miles, & Irons, 2004) and although the self-compassion subscale was positively correlated with the SCS (r = .65), this is arguably weaker than might be expected given they allegedly measure the same construct. Only one subgroup was analyzed for interpretability (Arts versus Engineering students), however the authors stated no predictions about differences between these groups.

Compassionate care assessment tool (CCAT; Burnell & Agan, 2013). The 28-item CCAT was developed to measure levels of compassion demonstrated by individual nurses providing care for patients in acute hospital settings. In contrast to the other scales reviewed so far, this scale is completed by patients in relation to their carers. Respondents rate compassionate care from two perspectives - the importance of each item to them personally, and the degree to which their individual nurses demonstrated these qualities. Ratings range from 1 to 4. A selection of possible items for the scale were derived from the Spiritual Needs Survey (Galek, Flannelly, Vane, & Galek, 2005) and the Caring Behaviors Inventory (Wu, Larrabee, & Putman, 2006), and refined after consulting with hospital staff involved in implementing national criteria for compassionate care, nurses, and patients. The CCAT focuses on four domains: the ability of carers to establish meaningful connection (e.g. having a sense of humor), to meet patient expectations (e.g. giving timely treatments), display caring

25

attributes (e.g. considering personal needs), and exhibit capable practitioner qualities (e.g. appearing competent).

Content validity. Overall, the scale was considered partially satisfactory for content validity. It includes items relating to three of the five elements in our definition of compassion: whether patients thought carers felt for them (emotional resonance), acted to help relieve their suffering, and could tolerate distress (e.g. asking if they 'remained calm at all times', treated them non-judgmentally, and excused their shortcomings). Items related to recognizing suffering and understanding the universality of suffering were not included. Additionally, some of the items are rather ambiguous - for example, one item asks whether nurses "addressed difficult issues", which could relate to their ability to tolerate distress, or to their ability to resolve more practical matters. It is also questionable whether the scale is actually measuring levels of compassion of nurses; factor analyses appear to have been carried out based on asking patients to rate how important each item was to them, rather than on asking them to rate the extent to which their carers behaved in this way.

Furthermore, as a number of items were derived from the Spiritual Needs Survey, there is a fairly strong emphasis on whether spiritual support was offered to patients, which is not necessarily relevant to the measurement of compassion for all patients. Similarly, several items taken from the Caring Behaviors Inventory ask about whether nurses gave timely treatments to patients, showed skill with equipment and helped control pain; however, while competence may be important in order to provide compassionate care, such abilities in themselves do not necessarily equate to showing compassion. It could also be argued that some of the areas tapped, such as providing timely treatments, controlling pain, and providing access to spiritual support, depend on variables outside of nurses' power (i.e. managerial or organizational level factors), and this raises a wider issue around the extent to which compassion can and should be measured at an individual or organizational level.

26

Factor structure and reliability. EFA supported a four-factor structure, with the four aforementioned domains. However, as previously noted, analyses appear to have been carried out based on asking patients to rate how important each item was to them, rather than on asking them to rate the extent to which their carers behaved in this way. This means that it is not clear whether the scale is measuring actual levels of compassion of their nurses per se. Additionally, the authors report that only 20 of the 28 items fit into the four factors identified, but they nonetheless appear to have retained all 28 items. Therefore, the CCAT was given a rating of 0 for factor structure. Cronbach's alpha values were adequate for the total scale and subscales. Test-retest reliability was not reported.

Convergent validity and interpretability. Convergent validity was not reported. Limited subgroup analyses were conducted for interpretability.

The Schwartz Center compassionate care scale (SCCCS; Lown, Muncer, & Chadwick, 2015). The 12-item SCCCS was developed to measure patients' ratings of compassionate inpatient care received from physicians' during a recent hospitalization. Patients complete items using a ten-point scale from 1 (not at all successful) to 10 (very successful). Items were initially developed by a committee consisting of patients, family members of patients, and individuals working in healthcare policy and advocacy, and were fine-tuned in five focus groups with patients, physicians, and nurses.

Content validity. Overall, the SCCCS was considered partially satisfactory for content validity. It includes items which could be interpreted to relate to three of the five elements in our definition of compassion: Whether patients thought physicians expressed sensitivity, care, and compassion for them (emotional resonance/acting to alleviate suffering), listened attentively (recognizing suffering), and acted in ways to relieve their suffering. The SCCCS did not appear to contain items related to understanding the universality of suffering and

tolerating uncomfortable feelings. Additionally, a couple of items refer to competence in caring (whether physicians spend enough time with patients, whether physicians communicate test results in a timely manner) which does not necessarily equate to showing compassion and could be dependent on factors outside of physicians' power (i.e. managerial or organizational level factors).

Factor structure and reliability. The SCCCS originally consisted of 16 items which were split into two item sets and administered to 801 recently hospitalized patients; half were asked item set one and half item set two. The authors conducted an EFA and CFA for each set of items and concluded that items within each set were unidimensional. However, they did not conduct analyses on all of the items, making it impossible to determine whether the scale as a whole is unidimensional, or whether the measure consisted of two separate scales or subscales. Despite this, the SCCCS was presented as a single scale. Although Cronbach's alpha values were adequate for both sets, these values were based on there being eight items in each set; the final 12-item scale consisted of seven items from the first set and five items from the second set after the removal of problematic items (e.g. items with lowest item-total correlations). The alpha value for all of the scale items was also missing. Test-retest reliability was not tested.

Convergent validity and interpretability. The authors found moderate to large, positive correlations between scores on both sets of items from the 12-item SCCCS and related constructs. Interpretability was not tested.

Discussion

The first aim of this paper was to synthesize existing conceptualizations of compassion and to propose a new definition that integrates common elements. A range of definitions from Buddhist and Western psychological perspectives were considered and five

components of compassion were identified: recognition of suffering; understanding its universality; feeling sympathy, empathy, or concern for those who are suffering (which we describe as emotional resonance); tolerating the distress associated with the witnessing of suffering; and motivation to act or acting to alleviate the suffering. Each of these components has been articulated by several published definitions of compassion, although no single existing definition explicitly includes all five of them. We do not claim that these five elements constitute statistically distinct factors of an overarching construct of compassion; this possibility must be empirically tested. However, we argue that our definition provides a useful foundation for the development of a comprehensive new measure of compassion.

The need for a new measure is supported by the findings of our review of existing measures of compassion. The maximum quality rating of any measure was seven out of a possible fourteen, suggesting that no scale exists that comprehensively measures compassion and provides scores with acceptable levels of reliability and validity. In other words, we cannot be confident that existing measures of compassion are measuring this construct accurately and this raises significant barriers to scientific progress in the field – how can we assess compassion and evaluate the effectiveness of interventions intended to enhance compassion if we cannot measure the construct accurately?

Quality ratings were low both because of poor ratings for content validity (the extent to which items appeared to fit our definition of compassion) and because of poor or untested psychometric properties. Internal consistency was strong for total scores but weak for many subscales. Evidence for the proposed factor structure of some scales was weak or absent. The presence of floor or ceiling effects was not examined for any scale, and test-retest reliability was examined for only one. Convergent correlations were generally significant and in the expected directions, but discriminant validity was not assessed. Low quality ratings could also be attributed to measures being in their early stages of development and initial papers

29

being unlikely to include a thorough test of psychometric properties. Quality ratings for compassion measures may improve over time with additional research including psychometric research.

The strongest measures identified were Neff's (2003b) Self-Compassion Scale and Hacker's (2008) Relational Compassion Scale, but neither of these measures capture each of the five elements in our definition. As Neff's measure focuses on self-compassion rather than compassion more generally or compassion for others, it is in any case not entirely suitable as a measure that can be used to determine levels of compassion in populations for whom compassion toward others is of interest (e.g., healthcare professionals). Given the current enthusiasm for compassion across different contexts, it is critical for future research to develop a psychometrically robust measure of the proposed definition of compassion as well as to explore more fully the relationship between self- and other-compassion.

Strengths and Limitations

A strength of this review is its contribution to greater clarity in the conceptualization of compassion and its components, which have previously been described in a variety of ways. The five elements of compassion extracted from our synthesis of definitions suggests that compassion is a complex construct that includes emotion but is more than an emotion, as it also includes perceptiveness or sensitivity to suffering, understanding of its universality, acceptance, nonjudgment, and distress tolerance, and intentions to act in helpful ways. This conceptualization suggests that compassion can be state-like and trait-like. Sensitivity to one's own or others' suffering, emotional responsiveness, acceptance and nonjudgment in the face of suffering, and motivation to be helpful are all likely to fluctuate across time and situations. On the other hand, Goetz et al (2010) present evidence suggesting that compassion can be seen as a trait-like quality that endures over time (e.g., Eisenberg et al., 2002). An

30

implicit assumption of compassion-focused interventions seems to be that a trait-like general tendency to be compassionate toward oneself or others can be developed through repeated practice of skills that cultivate compassionate states, attitudes, or behaviours. Additionally, although many of the questionnaires reviewed treat compassion as a disposition that is fairly consistent across contexts, some measures conceptualize compassion as operating within a particular context or social interaction (e.g., the CCAT).

This review assumes that compassion can indeed be measured with questionnaire methods. Some authors have suggested that subtle but observable behaviours, such as using a soft tone of voice, may also be valid indicators of compassion (Cameron, Mazer, DeLuca, Mohile, & Epstein, 2013), while Pearson (2006) notes that acts of compassion are often 'invisible', being "simple not clever; basic not exquisite; peripheral not central" (p. 22). This means that, as Dewar, Pullin, and Tocheris (2011) note, "there is a danger, therefore, of measuring what is easy to quantify, rather than what is important" (p. 32). Dewar et al. also point out that compassion can be seen as something that is negotiated between individuals in their interactions. These points suggest that, as with many psychological variables, questionnaire measures may only provide a partial picture of compassion. Furthermore, while questionnaire measures benefit from being simple to administer and complete, and helpful for tapping people's underlying attitudes where these are not directly observable, it may be difficult for people to complete such measures accurately in some contexts, for example in situations where healthcare staff feel under threat to be seen to be compassionate.

A further limitation of this review is the approach taken to identifying the definitions of compassion in Table 1. A systematic search was attempted but the way the field has evolved does not easily lend itself to a systematic review. Defining compassion was very rarely the primary purpose of papers; definitions were typically embedded as secondary to addressing the primary purpose of the paper. An early search generated an unfeasibly large

number of results. We therefore relied on the expertise of the authors to identify key theorists and sources in the field.

The review also assumed that individual levels of compassion should be measured. However, it has been argued that measuring compassion at the individual level opens people to accusations that they are not sufficiently compassionate. For example, in a healthcare context, this may result in a tendency to blame healthcare professionals for failings that in fact relate to external factors such as resourcing pressures or organizational restructuring (Crawford et al., 2014). This is an important consideration and highlights the need to ensure that efforts to measure levels of compassion among individuals do not serve to overstate individual deficits while deflecting attention from the broader impact of resourcing constraints and wider organizational changes.

Future Research

This review has argued that currently no psychometrically robust self- or observer-rated measure of compassion exists, despite widespread interest in measuring and enhancing compassion towards self and others. Future research should therefore focus on developing a psychometrically robust questionnaire-based measure of compassion, while keeping in mind the complexities around measuring this construct. It will subsequently be of value for future research to identify interventions (at both an individual level and organizational level) that have the potential to enhance compassion and examine whether changes in compassion mediate the outcomes of these interventions.

Although our review provides a foundation for progressing research into compassion, it represents a starting point. Future work should articulate theory driven hypotheses that test the relationships between key constructs and the validity of our five-element definition of compassion. This will generate important new knowledge about how these different

32

elements interact to give rise to compassion. It may be that some elements are facilitators of compassion or emergent factors rather than defining features.

Using a range of designs (including prospective and experimental designs), and triangulation of measurement to include behavioral (e.g., observable compassionate responses), bio-behavioral measures (e.g., as derived from Gilbert's theory) and self-report measures, will further aid the development of theory and understanding. It is likely that this will have real practical implications for how best to cultivate compassion in ways that support resilience and well being at both personal and societal levels.

Conclusion

In recent years, compassion has received increased scientific interest. Compassion has been defined here, in line with the literature, as involving five elements: recognizing suffering in others; understanding the common humanity of this suffering; feeling emotionally connected with the person who is suffering; tolerating difficult feelings that may arise; and acting or being motivated to act to help the person. A systematic search of measures of compassion was undertaken but all of the identified measures were found to have notable psychometric weaknesses. This is a serious limitation in the field. For example, without adequate measures, we cannot determine with any confidence levels of compassion or whether interventions designed to enhance compassion are effective. Therefore, we now call for empirical testing of our five element definition and the development of a measure of compassion, following good practice guidelines to identify items and to test its psychometric properties.

References

American Medical Association. (2001). Principles of Medical Ethics. Retrieved from

http://www.ama-assn.org/ama/pub/physician-resources/medical-ethics/code-medical-ethics/principles-medical-ethics.page

Barker, C., Pistrang, N., & Elliott, R. (2002). Research methods in clinical psychology. London: Wiley.

Burnell, L., & Agan, D. L. (2013). Compassionate care: Can it be defined and measured? The development of the Compassionate Care Assessment Tool. International Journal of Caring Sciences, 6(2), 180-187.

Cameron, R. A., Mazer, B. L., DeLuca, J. M., Mohile, S. G., & Epstein, R. M. (2015). In search of compassion: A new taxonomy of compassionate physician behaviours. Health Expectations: An International Journal of Public Participation in Health Care and Health Policy, 18(5), 1672-1685.

Cassell, E. J. (2002). Compassion. In C. R. Synder & S. J. Lopez (Eds.), Handbook of positive psychology (pp. 434-445). Oxford: Oxford University Press.

Chadwick, P., Hember, M., Symes, J., Peters, E., Kuipers, E., & Dagnan, D. (2008). Responding mindfully to unpleasant thoughts and images: Reliability and validity of the southampton mindfulness questionnaire (SMQ). British Journal of Clinical Psychology, 47(4), 451-455.

CoED Foundation. (2014). Bringing compassion into education and learning: The second strategic plan. Retrieved from http://www.coedfoundation.org.uk/pdfs/CoEd_Foundation_strategic_plan.pdf

Cosley, B. J., McCoy, S. K., Saslow, L. R., & Epel, E. S. (2010). Is compassion for others stress buffering? Consequences of compassion and social support for physiological reactivity to stress. Journal of Experimental Social Psychology, 46, 816–823.

Costa, J., Marôco, J., Pinto-Gouveia, J., Ferreira, C., & Castilho, P. (2015). 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.

Crawford, P., Brown, B., Kvangarsnes, M., & Gilbert, P. (2014). The design of

compassionate care. Journal of Clinical Nursing, 23(23-24), 3589-3599.

Dalai Lama. (1995). The power of compassion. New Delhi: HarperCollins.

Dalai Lama. (2002). How to practice: The way to a meaningful life. New York: Pocket Books.

Dalai Lama. (2005). Essence of the Heart Sutra: The Dalai Lama's Heart of Wisdom

Teachings. USA: Wisdom Publications.

Darwin C. (1871). The descent of man, and selection in relation to sex. London: John Murray.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a

multidimensional approach. Journal of Personality and Social Psychology, 44, 113-126.

de Waal, F. (2009). The age of empathy. New York: Harmony.

Department of Health. (2013). The handbook to the NHS constitution. Retrieved from

https://www.gov.uk/government/publications/the-nhs-constitution-for-england

Dewar, B., Pullin, S., & Tocheris, R. (2011). Valuing compassion through definition and

measurement. Nursing Management, 17(9), 32-7.

Eisenberg, N., Guthrie, I. K., Cumberland, A., Murphy, B. C., Shepard, S. A., Zhou, Q., &

Carlo, G. (2002). Prosocial development in early adulthood: A longitudinal

study. Journal of Personality and Social Psychology, 82(6), 993-1006.

Eisenberg, N., VanSchyndel, S. K., & Hofer, C. (2015). The association of maternal

socialization in childhood and adolescence with adult offsprings' sympathy/caring.

Developmental Psychology, 51(1), 7-16.

Feldman, C., & Kuyken, W. (2011). Compassion in the landscape of suffering.

Contemporary Buddhism, 12(1), 143-155.

Galek, K., Flannelly, K. J., Vane, A., & Galek, R. M. (2005). Assessing a patient's spiritual needs: A comprehensive instrument. Holistic Nursing Practice, 19(2), 62-69.

Gilbert, P. (2005). Compassion and cruelty: A biopsychosocial approach. In P. Gilbert (Eds.), Compassion: Conceptualisations, research and use in psychotherapy (pp. 9–74). London: Routledge.

Gilbert, P. (2009). *The Compassionate Mind: A new approach to life's challenges*. London: Constable and Robinson.

Gilbert, P. (2010). The compassionate mind. London: Constable & Robinson Ltd.

Gilbert, P., Clarke, M., Hempel, S., Miles, J. N. V., & Irons, C. (2004). Criticizing and reassuring oneself: An exploration of forms, styles and reasons in female students. British Journal of Clinical Psychology, 43(1), 31-50.

Gillath, O., Shaver, P. R., & Mikulincer, M. (2005). An attachment-theoretical approach to compassion and altruism. In In P. Gilbert (Eds.), Compassion: Conceptualisations, research and use in psychotherapy (pp. 122–147). London: Routledge.

Goetz, J. L., Keltner, D., & Simon-Thomas, E. (2010). Compassion: An evolutionary analysis and empirical review. Psychological Bulletin, 136(3), 351-374.

Hacker, T. (2008). The relational compassion scale: Development and validation of a new self-rated scale for the assessment of self-other compassion (Doctoral dissertation, University of Glasgow).

Hogan, R. (1969). Development of an empathy scale. Journal of Consulting and Clinical Psychology, 33(3), 307-316.

Hwang, J. Y., Plante, T., & Lackey, K. (2008). The development of the Santa Clara brief compassion scale: An abbreviation of Sprecher and Fehr's compassionate love scale. Pastoral Psychology, 56(4), 421-428.

Kanov, J. M., Maitlis, S., Worline, M. C., Dutton, J. E., Frost, P. J., & Lilius, J. M.

(2006). Compassion in organizational life Jossey-Bass, San Francisco, CA.

Keltner, D. (2009). Born to be good: The science of a meaningful life. New York: Norton.

Klimecki, O. M., Leiberg, S., Ricard, M., & Singer, T. (2014). Differential pattern of

functional brain plasticity after compassion and empathy training. Social Cognitive

and Affective Neuroscience, 9(6), 873-879.

Lazarus, R. S. (1991). Emotion and adaptation. Oxford: Oxford University Press.

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), e0132940.

Lown, B. A., Muncer, S. J., & Chadwick, R. (2015). Can compassionate healthcare be

measured? The Schwartz Center Compassionate Care Scale™. Patient education and

counselling, 98, 1005-1010.

MacBeth, A. & Gumley, A. (2012). Exploring compassion: a meta-analysis of the association

between self-compassion and psychopathology. Clinical Psychology Review, 32, 545-552

Martins, D., Nicholas, N. A., Shaheen, M., Jones, L., & Norris, K. (2013). The development

and evaluation of a compassion scale. Journal of Health Care for the Poor and

Underserved, 24(3), 1235-1246.

Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. Journal of

Personality, 40(4), 525-543.

Najjar, N., Davis, L. W., Beck-Coon, K., & Carney Doebbeling, C. (2009). Compassion

fatigue: a review of the research to date and relevance to cancer-care providers.

Journal of Health Psychology, 14, 267–77.

Neff, K. D. (2003a). Self-compassion: An alternative conceptualization of a healthy attitude

toward oneself. Self and Identity, 2(2), 85-101.

Neff, K. D. (2003b). The development and validation of a scale to measure self-compassion.

Self and Identity, 2(3), 223–250.

Neff, K. D. (2015). The Self-Compassion Scale is a Valid and Theoretically Coherent

Measure of Self-Compassion. Mindfulness, 1-11.

Neff, K. D., & Pommier, E. (2013). The relationship between self-compassion and other-

focused concern among college undergraduates, community adults, and practicing

meditators. Self and Identity, 12(2), 160-176.

Norko, M. A. (2005). Commentary: compassion at the core of forensic ethics. Journal of the

American Academy of Psychiatry and the Law, 33, 386-9

Pearson, A. (2006). Powerful caring. Nursing Standard (Royal College of Nursing (Great

Britain): 1987), 20(48), 20-22.

Pommier, E. A. (2010). The compassion scale (Order No. 3445994). Available from

ProQuest Dissertations & Theses Global. (855633530).

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 &

Psychotherapy, 18(3), 250-255.

Rosenberg, M. (1965). Society and the adolescent self-image. Princeton, NJ: Princeton UP

Sprecher, S., & Fehr, B. (2005). Compassionate love for close others and humanity. Journal

of Social and Personal Relationships, 22(5), 629-651.

Terwee, C. B., Bot, S. D. M., de Boer, M.,R., van der Windt, D.,A.W.M., Knol, D. L.,

Dekker, J., . . . de Vet, H.,C.W. (2007). Quality criteria were proposed for

measurement properties of health status questionnaires. Journal of Clinical

Epidemiology, 60(1), 34-42.

Wang, S. (2005). A conceptual framework for integrating research related to the physiology

of compassion and the wisdom of Buddhist teachings. In P. Gilbert (Eds.),

Compassion: Conceptualisations, research and use in psychotherapy (pp. 75–120).

London: Routledge.

Williams, M. J., Dalgleish, T., Karl, A., & Kuyken, W. (2014). Examining the factor structures of the five facet mindfulness questionnaire and the self-compassion scale. Psychological Assessment, 26(2), 407-418.

Wispe, L. (1991). The psychology of sympathy. New York: Plenum.

Wu, Y., Larrabee, J. H., & Putman, H. P. (2006). Caring behaviors inventory: A reduction of the 42-item instrument. Nursing Research, 55(1), 18-25.
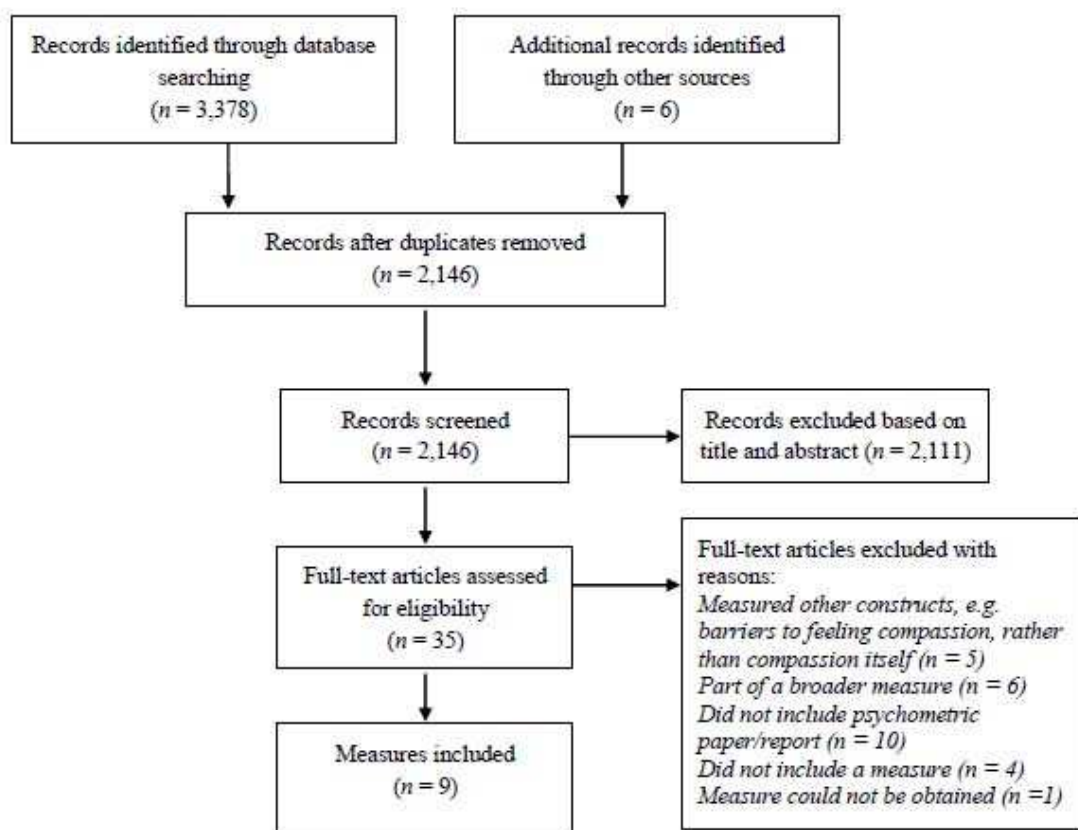
Figure 1. PRISMA flow diagram of search strategy.

Table 1. Major definitions of compassion in the literature in relation to the five-element definition of compassion

Major definitions of compassion in the literature in relation to the five-element definition of compassion Definition of Compassion	Recognizing suffering [a]	Understanding the universality of suffering	Emotional resonance	Tolerating uncomfortable feelings	Motivation to act/acting to alleviate suffering
1. "Being moved by another's suffering and wanting to help" (Lazarus, 1991, p.289).	✓(implied)		✓		✓
2. An openness to the suffering of others with a commitment to relieve it (Dalai Lama, 1995). Buddhist conceptualizations also highlight cognitive components (e.g. the ability to imagine and reason about a person's experiences) and approaching those who are suffering with tolerance and non-judgement.	✓(implied)	✓	✓	✓	✓
3. "Being touched by the suffering	✓ (explicitly stated)		✓	✓	✓

of others, opening one's awareness to others' pain and not avoiding or disconnecting from it, so that feelings of kindness towards others and the					
4. Compassion consists of three facets: Noticing, feeling, and responding (Kanov et al., 2004).	✓ (explicitly stated)		✓		✓
5. "A deep awareness of the suffering of another coupled with the wish to relieve it" (Gilbert, 2009, p. 13). Compassion consists of six 'attributes': Sensitivity, Sympathy, Empathy, Motivation/Caring, Distress Tolerance, and Non-Judgement.	✓ (explicitly stated)		✓	✓	✓
6. "The feeling that arises in witnessing another's suffering and that motivates a subsequent desire to help" (Goetz	✓ (explicitly stated)		✓		✓

et al., 2010, p.351).					
7. "An orientation of mind that recognises pain and the universality of pain in human experience and the capacity to meet that pain with kindness, empathy, equanimity and patience" (Feldman & Kuyken, 2011, p. 145).	✓ (explicitly stated)	✓	✓	✓	
8. Compassion involves three elements: Kindness, mindfulness, and common humanity (Pommier, 2011).	✓(implied)	✓	✓	✓	✓

a Some definitions of compassion explicitly include an element of 'recognizing suffering', whereas in others, this is implied. We have indicated whether 'recognizing suffering' is explicitly stated or implied in the following way: ✓(explicitly stated) and ✓(implied).

Table 2. Quality ratings of measures of compassion

Measure	Content validity	Factor structure	Internal consist-ency	Test retest reliability	Convergent / discriminant validity	Floor/ ceiling effects	Interpret-ability	**Total /14**
SCS	1	0	2	2	1	0	1	**7**
RCS	1	2	2	0	1	0	1	**7**

CLS	1	1	2	0	1	0	1	**6**
SCBCS	1	1	2	0	1	0	1	**6**
CS-P	1	2	1	0	1	0	1	**6**
CS-M	1	0	2	0	0	0	1	**4**
CCAT	1	0	2	0	0	0	1	**4**
SCS-SF	1	2	1	0	0	0	0	**4**
SCCCS	1	0	0	0	1	0	0	**2**

Note. Rating: 0 = criterion not met/insufficient data to rate criterion; 1 = criterion partially

met; 2 = criterion fully met.

CCAT = Compassionate Care Assessment Tool; CLS = Compassionate Love Scale; CS-P = Pommier Compassion Scale; CS-M = Martins et al. Compassion Scale; RCS = Relational Compassion Scale; SCBCS = Santa Clara Brief Compassion Scale; SCCCS = Schwartz Center Compassionate Care Scale; SCS = Self-Compassion Scale; SCS-SF = Self-Compassion Scale – Short Form.

Table 3. Psychometric properties of measures of compassion (content validity, factor

structure, internal consistency, and test-retest reliability).

Measure	Content validity: Factors captured[a] /5	Content validity: Item generation (recipient and	Proposed factor structure	Support for factor structure: Type of analysis conducted (factor	Internal consisten cy: Adequate sample size for factor	Internal consisten cy: Cronbach alpha (for total scale and	Test retest reliabilit y: r (time between testing)

		expert groups consulted?)	structure found)[b]		analyses?		subscales)
CLS	4 (U, ER, T, & A)	Recipients = yes Experts = yes	Not reported	EFA (single factor for both versions)	Yes (N = 354)	$\alpha = .95$ for both close others and strangers-humanity versions	Not reported
SCBCS	2 (ER & A)	Recipients = yes Experts = yes	Not reported	EFA (single factor)	Yes (N = 223)	$\alpha = .90$	Not reported
CS-M	1 (A)	Recipients = yes Experts = yes	Five factors	EFA (two-factor structure found but rejected in favour	Yes (N = 310)	$\alpha = .82$	Not reported

				of a single-factor model)			
SCS	4 (U, ER, T, & A)	Recipients = no Experts = yes	Six factors represented under a single overarching construct	CFA (six factors)	Yes (N = 391)	Total α = .92, Subscales = .75 to .81.	Total scale: r = .93, Subscales: r = .80 to .88 (3 weeks)
SCS-SF	4 (U, ER, T, & A)	Recipients = no Experts = yes	Six factors represented under a single overarching construct	CFA (six factors represented under a single overarching construct)	Yes (N = 415)	Total α = .86, Subscales = .54 to .75.	Not reported

CS-P	4 (R, U, ER, & A)	Recipients = no Experts = yes	Six factors represented under a single overarching construct	CFA (six factors represented under a single overarching construct)	Yes (sample 1: N = 439, sample 2: N = 510)	Total α = .90 (sample 1) and .87 (sample 2). Subscale αs < .70 for 4/6 subscales in sample 1 and 1/6 subscales in sample 2.	Not reported
RCS (compassion for others and self-compassion subscales)	4 (R, ER, T, & A)	Recipients = no Experts = yes	Four factors	CFA (four factors)	Yes (N = 231)	Subscales = .74 to .84	Not reported

47

CCAT	3 (ER, T, & A)	Recipients = yes Experts = yes	Not reported	EFA (four factors)	Yes (N = 250)	Total α > .70 (exact value not given), Subscales = .77 to .87.	Not reported
SCCCS	3 (R, ER, A)	Recipients = yes Experts = yes	Not reported	EFA (single factor but analysis was not conducted on all items) CFA (single factor but analysis was not conducted on all items)	Yes (N = 801)	Subscales = .97 and .95. Single scale but total α missing.	Not reported

48

CCAT = Compassionate Care Assessment Tool; CFA = confirmatory factor analysis; CLS = Compassionate Love Scale; CS-P = Pommier Compassion Scale; CS-M = Martins et al. Compassion Scale; EFA = exploratory factor analysis; RCS = Relational Compassion Scale; SCBCS = Santa Clara Brief Compassion Scale; SCCCS = Schwartz Center Compassionate Care Scale; SCS = Self-Compassion Scale; SCS-SF = Self-Compassion Scale – Short Form.

[a]Five elements of compassion: R = recognising suffering; U = understanding the universality of suffering; ER = emotional resonance; T = tolerating uncomfortable feelings; A = acting or motivation to act to alleviate suffering.

[b]For details of the factors identified, refer to the results section.

Table 4. Psychometric properties of measures of compassion (convergent validity and interpretability)

Measure	Convergent validity: Correlation (Pearson's r) of compassion measure with measures of related constructs	Interpretability: Subgroups tested for differences
CLS	PSP other-oriented empathy subscale and empathy items from Schieman & Van Gundy (2000): r = .50 to .68; PSP helpfulness subscale: r = .23 (close others), r = .32 (strangers); Frequency of church attendance: r = .22 (close others), r = .43 (strangers); Volunteerism items from Mikulincer et al. (2005): r = .18 (close others), r = .35	Gender (women scored significantly higher than men on both versions)

	(strangers);	
	Social support (developed by authors): r = .27 (strangers), r = .51 (close others);	
	DSES: r = .39 (close others), r = .44 (strangers).	
SCBCS	IRI empathic concern subscale: r = .65; VIQ: r = .48; SCSORF: r = .27.	Gender (women scored significantly higher than men)
CS-M	CLS: r = .66	Gender (women scored significantly higher than men), Education (significantly higher for those with University education versus High School or less), Income (significantly higher for those with annual income $40,000+ versus $10,000 or less), Age, race, and marital status (no differences)
SCS	DEQ self-criticism subscale: r = -.65; SOC: r = .41; TMMS attention subscale: r = .11; TMMS clarity subscale: r = .43; TMMS repair subscale: r = .55; RSES: r =	Gender (women scored significantly lower than men)

50

.59.

SCS-SF	Not reported	Not reported
CS-P	SCS: r = .01; SOC: r = .41; 3D-WS reflective subscale: r = .26; 3D-WS cognitive subscale: r = .39; 3D-WS affective subscale: r = .56; QMEE: r = .58; IRI empathic concern subscale: r = .65; IRI perspective taking subscale: r = .35; CLS close others version: r = .54; CLS strangers version: r = .27; SMQ: r = -.12.	Gender (women scored significantly higher than men)
RCS (compassion for others and self-compassion subscales)	RCS compassion for others subscale & SCS: r = .24; RCS self-compassion subscale & SCS: r = .65; RCS compassion for others subscale & EACS emotional expression & processing subscales: r = .41 and .42; RCS self-compassion subscale & EACS emotional expression & processing subscales: r = .51 and .46; RCS compassion for others subscale & SCSRS inadequate & hated self subscales: r = .03 and .12;	Significant differences in RCS scores between Arts and Engineering students. The direction of the results for each subscale was not specified.

RCS self-compassion subscale & SCRS

inadequate & hated self subscales:

r = -.29 and -.41;

RCS compassion for others subscale &

SCSRS reassured self subscale: r = .01;

RCS self-compassion subscale & SCRS

reassured self subscale: r = .43;

RCS compassion for others subscale and

RSQ secure attachment: r = .34;

RCS self-compassion subscale and RSQ

secure attachment: r = .31;

RCS compassion for others subscale and

RSQ insecure attachment styles (fearful,

preoccupied, dismissing, anxious, and

avoidant):

r = -.23, -.06, -.15, -.19, and -.22,

respectively;

RCS self-compassion subscale and RSQ

insecure attachment styles (fearful,

preoccupied, dismissing, anxious, and

avoidant):

r = -.22, -.15, -.05, -.03, and -.07,

respectively.

| CCAT | Not reported | Gender (women scored carers |

		significantly higher than men),
		Marital status and reason for
		hospitalisation (no differences)
SCCCS	Overall satisfaction with recent hospitalisation (item set 1): r = .54; Overall satisfaction with recent hospitalisation (item set 2): r = .60; Satisfaction with communication and emotional support (item set 1): r = .72; Satisfaction with communication and emotional support (item set 2): r = .64.	Not reported

3D-WS = 3-Dimensional Wisdom Scale (Ardelt, 2003); CCAT = Compassionate Care Assessment Tool; CLS = Compassionate Love Scale; CS-P = Pommier Compassion Scale; CS-M = Martins et al. Compassion Scale; DEQ = Depressive Experiences Questionnaire (Blatt, D'Afflitti, & Quinlan, 1976); DSES = Daily Spiritual Experience Scale (Underwood & Teresi, 2002); EACS = Emotional Approach Coping Scale (Stanton et al., 2000); IRI = Interpersonal Reactivity Index (Davis, 1980); PSP = Prosocial Personality Battery (Penner et al., 1995); QMEE = Questionnaire Measure of Emotional Empathy (Mehrabian & Epstein, 1972); RCS = Relational Compassion Scale; RSES = Rosenberg Self-Esteem Scale (Rosenberg, 1965); RSQ = Relationship Scales Questionnaire (Griffin & Bartholomew, 1994); SCBCS = Santa Clara Brief Compassion Scale; SCSORF = Santa Clara Strength of Religious Faith Questionnaire (Plante & Boccaccini, 1997); SCCCS = Schwartz Center Compassionate Care Scale; SCS = Self-Compassion Scale; SCS-SF = Self-Compassion Scale – Short Form; SCSRS = Self-Criticising/Attacking and Self-Reassuring Scale (Gilbert et al., 2004); SMQ = Southampton Mindfulness Questionnaire (Chadwick et al., 2008); SOC =

Social Connectedness Scale (Lee & Robbins, 1995); TMMS = Trait Meta-Mood Scale

(Salovey et al., 1995); VIQ = Vocational Identity Questionnaire (Dreher et al., 2007).

[a] Five elements of compassion: R = recognising suffering; U = understanding the universality

of suffering; ER = emotional resonance; T = tolerating uncomfortable feelings; A = acting or

motivation to act to alleviate suffering.

Highlights

- Compassion is recognized as important across many sectors of society

- There is lack of consensus on definition and few self/observer-rated measures exist

- Five elements of compassion are proposed after consolidating existing definitions

- The psychometric properties of existing measures are poor, limiting their utility

- A new measure of compassion with robust psychometric properties is needed