

# Effectiveness of self-compassion-related interventions for reducing self-criticism: A systematic review and meta-analysis

Katherine E. Wakelin  | Gemma Perman  | Laura M. Simonds 

School of Psychology, University of Surrey,  
Guildford, UK

## Correspondence

Katherine Wakelin, School of Psychology,  
University of Surrey, Guildford GU2 7XH, UK.  
Email: k.wakelin@surrey.ac.uk

## Abstract

Self-criticism is the process of negative self-evaluation. High levels are associated with psychopathology and poorer therapeutic outcomes. Self-compassion interventions were developed to explicitly target self-criticism. The aim of this review was to estimate the overall effect of self-compassion-related interventions on self-criticism outcomes and investigate potential moderating variables. A systematic search of the literature identified 20 randomized controlled trials (RCTs) that met the inclusion criteria. Nineteen papers, involving 1350 participants, had sufficient data to be included in the meta-analysis. Pre- and post-data points were extracted for the compassion and control groups. Study quality was assessed using an adapted version of the Cochrane Collaboration's risk of bias tool, which concluded that studies were of moderate quality. Meta-analysis findings indicated that self-compassion-related interventions produce a significant, medium reduction in self-criticism in comparison with control groups (Hedges'  $g = 0.51$ , 95% CI [0.33–0.69]). Moderator analysis found greater reductions in self-criticism when self-compassion-related interventions were longer and compared with passive controls rather than active. The remaining moderators of forms of self-criticism, sample type, intervention delivery, intervention setting and risk of bias ratings were insignificant. Overall, the review provides promising evidence of the effectiveness of self-compassion-related interventions for reducing self-criticism. However, results are limited by moderate quality studies with high heterogeneity. Directions for future research indicate that more RCTs with active controls, follow-ups, consistent use and reporting of measures and diverse samples are needed.

## KEYWORDS

intervention, meta-analysis, randomized controlled trials, review, self-compassion, self-criticism

## 1 | INTRODUCTION

Self-criticism is the process of negative self-evaluation and self-scrutiny (Kannan & Levitt, 2013; Shahar, 2015), which is accompanied by negative emotions such as anger and self-contempt (Whelton &

Greenberg, 2005). Early cognitive-based theories tended to conceptualize self-criticism as a single process varying in degree of severity (Beck et al., 1979; Blatt et al., 1982). However, subsequent research by Gilbert and others propose that there are two different forms of self-criticism, referred to as the 'hated-self' and the 'inadequate self',

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Clinical Psychology & Psychotherapy* published by John Wiley & Sons Ltd.

which each serve a unique function and have a variable impact on psychological distress (Gilbert et al., 2004; Halamová, Kanovský, Gilbert, et al., 2018). Gilbert et al.'s (2004) two forms of self-criticism have been replicated in different demographics, including in Portuguese samples (Castilho et al., 2015; Kupeli et al., 2013), inpatients and day-patients from mixed clinical populations (Gilbert et al., 2010) and patients diagnosed with schizophrenia (Mayhew & Gilbert, 2008).

The 'hated-self' form of self-criticism is focused on aggressive, self-hate and a desire to 'get rid' of perceived undesirable aspects of the self with the function of self-persecution. The 'inadequate-self' is focused on drawing attention to inadequacies, things to improve and failures, with the function of self-correction. In support of the argument that self-criticism has discrete functions, neural research has found that the inadequate-self form is related to reassuring processes in the brain, which is not the case for the hated-self form (Kim, Henderson, et al., 2020). This suggests that there is an aspect of the inadequate-self that is more positively orientated towards self-reassurance when self-hate is controlled for (Gilbert et al., 2004). Therefore, it is unsurprising that the hated-self form of self-criticism is more strongly associated with psychopathology than the inadequate-self form (Gilbert et al., 2004). However, both the hated-self and inadequate-self forms negatively correlate with self-esteem, life satisfaction and self-compassion and discriminate between clinical and nonclinical populations (Biermann et al., 2020; Castilho et al., 2015). Therefore, the different forms of self-criticism are important to consider in relation to psychological distress.

Although self-evaluation is a process inherent to human existence, individuals who experience high levels of self-criticism experience a hostile, self-berating internal monologue that is associated with feelings of unworthiness, failure, shame, guilt, paranoia and attachment difficulties (Blatt & Zuroff, 1992; Cantazaro & Wei, 2010; Gilbert & Miles, 2000). Over time, high levels of self-criticism are found to associate with self-injurious behaviour and suicidality (Fazaa & Page, 2009; Glassman et al., 2007; Klomek et al., 2008). Research from a diagnostic perspective reports a consistent association between high self-criticism and a range of psychopathology, including depression (Luyten et al., 2007), social anxiety (Cox, Fleet, & Stein, 2004), eating disorders (Dunkley & Grilo, 2007), borderline personality disorder (Kopala-Sibley et al., 2012), persecutory delusions (Hutton et al., 2013) and posttraumatic stress disorder (Cox, MacPherson, et al., 2004). The breadth of these findings indicates the transdiagnostic nature and clinical significance of self-criticism in psychological distress.

Individuals with higher levels of self-criticism tend to have poorer outcomes in traditional therapies, such as cognitive therapy and cognitive behavioural therapy (CBT) (Enns et al., 2002; Rector et al., 2000). It is thought that this is because self-critical individuals struggle to build interpersonal relationships and find it difficult to build a positive therapeutic alliance (Blatt et al., 1995; Shahar et al., 2003; Zuroff et al., 1999, 2000). This is understandable considering that high levels of self-criticism may be associated with inconsistent or unresponsive attention from early caregivers and the

### Key Practitioner Message

- The meta-analysis found that self-compassion-related interventions produce a medium reduction in self-criticism in comparison with control groups in a range of samples.
- Greater reductions in self-criticism were seen when self-compassion-related interventions were compared to passive controls, rather than active. However, more randomized controlled trials (RCTs) using active controls are recommended.
- Longer self-compassion-related interventions were associated with greater effect sizes.
- Moderator analyses of forms of self-criticism, sample type, intervention delivery, intervention setting and risk of bias ratings lacked power. Further studies are needed to assess if these factors have meaningful effects on the reduction of self-criticism in self-compassion-related interventions.

development of attachment anxiety or attachment avoidance (Bowlby, 1969, 1973; Cantazaro & Wei, 2010; Naismith et al., 2019). At a neurobiological level, research has found that when individuals engage in self-criticism, attachment processes modulate mental imagery and threat responses (Kim, Kent, et al., 2020). Kim, Kent, et al. (2020) suggest that when the threat response is activated during self-criticism, mental visual imagery may be higher amongst individuals with secure attachments and lower amongst those with avoidant attachments. This illustrates how attachment patterns may impact the psychological process of self-criticism. However, the degree to which self-criticism is successfully reduced in therapy has been found to be a stronger predictor of treatment response than pretherapy self-criticism levels (Rector et al., 2000). Therefore, self-criticism is an important target for interventions aiming to prevent or treat mental health difficulties given its association with psychological distress and its potential influence on therapeutic outcomes.

Compassion has received growing attention as the 'antidote' to self-criticism and means by which to counteract related distress (Longe et al., 2010; Lutz et al., 2020). This is because cultivating compassion towards the self enables individuals to develop a compassionate internal voice to counteract internal self-critical voices and reduce negative emotions (Gilbert, 2010). Various definitions of compassion have been proposed in the literature, and this has given rise to multiple forms of self-compassion-related interventions. One definition by Gilbert (2014, p. 19) is that compassion is 'the sensitivity to suffering in self and others (engagement), with a commitment to try and alleviate and prevent it (action)'. This has led to the development of compassion-focused therapy (CFT), informed by evolutionary psychology, attachment theory, Buddhist philosophy and social mentality theory (Gilbert, 2014). The aim of CFT is to move individuals from operating in a competitive, social rank motivation system to a

compassionate motivation system (Gilbert, 2014). CFT uses techniques such as imagery, meditation, letter writing and role play to explicitly develop courage, wisdom and self-soothing abilities in clinical populations with high levels of self-criticism and shame who do not respond well to traditional therapies (Gilbert, 2009; Gilbert & Procter, 2006).

Another perspective of compassion is Neff's (2003a, 2003b) conceptualization, also informed by Buddhist philosophy. It describes self-compassion as being made up of three components: mindfulness, rather than over identifying with difficulties; connection with others, rather than isolating oneself; and self-kindness, rather than being self-judgmental (Neff, 2003a). This school of thought has developed the Mindful Self-Compassion Programme, a structured programme aiming to develop self-compassion (Neff & Germer, 2013). However, unlike CFT, it targets nonclinical populations and has a heavier emphasis on cultivating mindfulness and appreciation. Many other self-compassion conceptualizations and interventions have been developed, which have also been influenced by Tibetan Buddhist traditions, such as Cultivating Emotional Balance (Kemeny et al., 2012), Loving-Kindness Meditations (LKMs) and Compassion Meditations (e.g., Wallmark et al., 2013).

Existing reviews indicate a growing evidence base for the effectiveness of self-compassion-related interventions for reducing anxiety, depression and psychological distress (Kirby et al., 2017), particularly for individuals with high levels of self-criticism (Leaviss & Uttley, 2015). Such effects are illustrated in studies using a range of clinical samples, such as eating disorders (Gale et al., 2014), personality disorders (Lucre & Corten, 2013), psychosis and paranoia (Braehler et al., 2013; Mayhew & Gilbert, 2008).

The current review takes a different focus to existing evidence reviews. Rather than excluding interventions because they are not explicitly CFT (Leaviss & Uttley, 2015) and are remote or only one session in length (Kirby et al., 2017; Wilson et al., 2019), any intervention describing content and techniques with the direct or indirect aim or expectation of cultivating a sense of self-compassion are included. In line with Kirby et al.'s (2017) recommendation, this allows a more comprehensive review investigating intervention differences as moderators to be conducted. Therefore, interventions more broadly conceptualizing self-compassion and its associated qualities, such as kindness and self-forgiveness, were included, such as LKMs. It is recognized that evidence by Gilbert et al. (2019) argue that kindness is distinct from compassion. However, as Gilbert et al.'s (2019) research was preliminary and this review does not define compassion solely in terms of Gilbert's (2010, 2014) CFT theory, LKMs were deemed appropriate to include. This decision was supported by the original objectives of LKM being not just to cultivate kindness but also to develop warmth and compassion (Luberto et al., 2018; Wallace, 1999), which evidence illustrates that LKMs successfully cultivate (Boellinghaus et al., 2014).

Second, samples with and without diagnosed mental health difficulties were included, unlike the recent **meta-analyses** that only considered participants with classifiable mental health symptoms (Craig et al., 2020; Wilson et al., 2019). This is because this review

recognizes self-criticism to be a pervasive construct on a continuum, impacting a variety of individuals across a range of settings and circumstances. The inclusion of nonclinical samples may introduce participants with lower levels of preintervention distress or less self-attacking forms of self-criticism into the review. This may reduce the potential intervention effect sizes as there may have been less potential for participant change. Therefore, sample type was investigated as a moderator to account for this.

Besides Kirby et al.'s (2019) meta-analysis examining fears of compassion, the majority of previous reviews have taken a diagnostic stance by analysing anxiety and depression outcomes (Ferrari et al., 2019; Kirby et al., 2017; Wilson et al., 2019). The authors are only aware of one review to date, Ferrari et al. (2019), which reviewed the effect of self-compassion on self-criticism. This is surprising considering that CFT was developed to explicitly target self-criticism. Ferrari et al.'s (2019) analysis was limited by only having eight randomized controlled trials (RCTs) available at the time. Thus, there is a clear rationale for an up-to-date review of the effect of self-compassion-related interventions on self-criticism, incorporating a broader range of interventions and samples than previous reviews.

The primary aim of this meta-analysis was to estimate the overall effect of self-compassion-related interventions on self-criticism, relative to control conditions in RCTs. The meta-analysis also sought to investigate the effect of intervention length, sample type, type of control comparison group, intervention delivery, intervention setting, and risk of bias on the relationship between therapy and self-criticism outcomes. Unlike any previous research known to the author, the paper also aims to investigate how different forms of self-criticism may affect the relationship between therapy and self-criticism outcomes.

## 2 | METHOD

### 2.1 | Search strategy

An electronic search was performed on the following databases to identify peer-reviewed RCTs published between January 1993 and January 2020: MEDLINE, PsycARTICLES, Psychology and Behavioural Sciences Collection and PsycINFO. The date limiter was chosen because the most evaluated compassion-based intervention, CFT, was developed by Paul Gilbert over the last 20 years. The following search terms were used: (1) compassion-focus\* OR CFT OR self-compass\* OR self-kindness OR "compassionate mind" OR CMT, (2) program\* OR training OR therapy OR intervention, (3) self-crit\* OR self-attack\* OR self-hat\* OR "self-directed hostility" OR self-judg\* OR self-condem\* OR "self-directed negative th\*". To improve the scope of the search, synonyms and alternative spellings were also included. Titles, abstracts and full texts were systematically searched using the inclusion and exclusion criteria outlined in Table 1 and the PRISMA guidelines (Moher et al., 2009).

A flow diagram of the process by which papers were identified and eliminated is included in Figure 1. A total of 1671 articles were

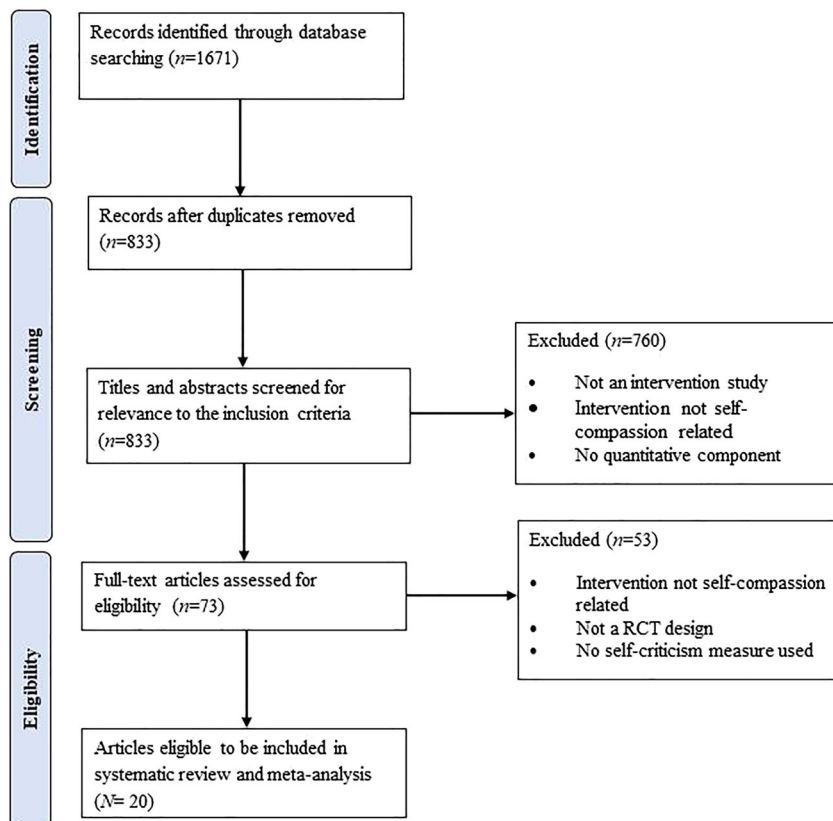
**TABLE 1** Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
<ul style="list-style-type: none"> <li>English language</li> <li>Published in a peer reviewed journal</li> <li>Published between Jan 1993 and Jan 2020</li> <li>Empirical study</li> <li>Has a RCT design (includes at least one active or passive control)</li> <li>Includes any intervention describing content and techniques that cultivate a sense of self-compassion. The ability to cultivate compassion will either be (1) an explicitly stated expectation or aim of the intervention or (2) be an indirect aim or expectation due to self-compassion being measured</li> <li>Includes a quantitative measure of self-criticism pre-post or post intervention (including measures of self-judgement, self-condemnation or self-directed negative thinking)</li> </ul>	<ul style="list-style-type: none"> <li>Not written in English</li> <li>Dissertations, book reviews or unpublished studies</li> <li>Published before Jan 1993</li> <li>Case studies, case reports, letters, books, reviews, editorials</li> <li>Non-RCT (i.e., cross sectional or within subjects)</li> <li>Intervention does not describe content and techniques that cultivate a sense of self-compassion</li> <li>Does not include a measure of self-criticism</li> </ul>

initially identified, 838 of which were removed due to duplicates. Using the inclusion and exclusion criteria (Table 1), the 833 papers remaining were then screened by their titles and abstracts, 73 of which were then reviewed as full-text. A sample of 20 papers was found to satisfy the inclusion and exclusion criteria.

## 2.2 | Study sample

Key information extracted from the papers about sample characteristics and study characteristics are represented in Tables 2 and 3, respectively. Halamová, Kanovský, Varšová, et al. (2018) study was the only paper not able to be included in the meta-analysis due to insufficient data reporting which was not able to be resolved by contacting the authors. Nevertheless, Halamová, Kanovský, Varšová, et al. (2018) are still reported in the tables of characteristics and quality assessment to enable a comprehensive review of the current literature. Table 2 highlights how the research was produced from a range of 12 countries, the majority being high-income countries using White Caucasian, American or Australian samples. All studies were published between 2013 and 2019. Across the 20 studies, 18 studies sampled predominantly, or entirely, female samples. Only six studies sampled individuals with diagnosed mental health difficulties and/or who were patients of mental health services (clinical samples). Thus, the majority of studies sampled nonclinical females from White, ethnic



**FIGURE 1** Adapted PRISMA flow-diagram illustrating the process of study selection [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

TABLE 2 Sample characteristics

Study	N recruited (Int/Con)	N analysed post intervention (Int/Con)	Attrition rate (% dropout)	Participant population	Recruitment sources	Gender (% female)	Mean age (SD)	Country	Ethnicity
Arimitsu, 2016	40 (20/20)	35 (19/16)	12.5%	University students with low self-compassion. Assessed: Scored below normative mean on SCS.	University website, clinics	Int: 85% Con: 65%	Int: 23.3 (7.4) Con: 19.4 (1.1)	Japan	Asian: 100%
Ascone et al., 2017	56 (unknown/Unknown)	51 (estimated: 26/25)	9%	Individuals from inpatient psychiatric wards or outpatients with a diagnosed psychotic disorder and paranoid delusions. Assessed: Structured DMS-IV clinical interviews or case-file ICD-10 diagnosis.	Psychiatric inpatient and outpatients	Int: 34% Con: 24%	Int: 40.2 (12.9) Con: 36.2 (10.1)	Germany	Not stated
Cornish & Wade, 2015	26 (15/11)	21(12/9)	19%	Individuals from the community experiencing shame over an offence committed against another person. Assessed: Clinical interview to exclude risk and psychotic disorders, criteria unclear.	Newspapers, flyers	77% Condition breakdown not given	36 (17) Condition breakdown not given	USA	European American: 80.8%, African American: 7.7%, Latina: 7.7%, Asian American: 3.8%
Duarte et al., 2017	33 (17/16)	20 (11/9)	39%	Individuals from the community with a diagnosed binge eating disorder. Assessed: diagnosis of being eating disorder, methodology unclear.	University, newspapers	100%	Int: 37.7 (7.5) Con: 35.8 (9.1)	Portugal	Caucasian: 100%
Dundas et al., 2017	138 (69/69)	117 (53/64)	15%	University students. Assessed: None.	University webinars	85% Condition breakdown not given	25 (4.9) Condition breakdown not given	Norway	Not stated
Fellu-Soler et al., 2017	32 (16/16)	32 (16/16) <sup>a</sup> Observed: Not stated	Unclear	Individuals from an outpatient psychiatric clinic with a diagnosis of borderline personality disorder. Assessed: DSM-IV diagnosis and structured diagnostic interview.	Outpatient clinic	Int: 100% Con: 88%	Int: 35.1 (8.3) Con: 32.5 (6.2)	Spain	Caucasian: 100%

(Continues)

TABLE 2 (Continued)

Study	N recruited (Int/Con)	N analysed post intervention (Int/Con)	Attrition rate (% dropout)	Participant population	Recruitment sources	Gender (% female)	Mean age (SD)	Country	Ethnicity
Halamová, Kanovský, Varšová, et al., 2018 <sup>b</sup>	123 (70/53)	55 (32/23)	55%	General community sample. Assessed: N/A	Social media, well-being forums	Int: 84% Con: 85%	Int: 33.7 (16.0) Con: 25.35 (6.3)	Slovakia	Caucasians and Slovaks: 100%
Johnson et al., 2018	59 (35/24)	Not stated	Unclear	Individuals who had recently attempted suicide and were patients in a public hospital providing care for low-income individuals. Assessed: Attempted suicide in the previous year and self-identified as African American.	Public hospital emergency rooms, inpatient units, outpatient clinics	Int: 54.3% Con: 50%	Int: 42.8 (9.0) Con: 44.4 (10.1)	USA	African American: 100%
Kelly & Carter, 2015	41 (15, 13, 13)	41 (15/13/13) <sup>a</sup> Observed: 35 (11/12/12) Meta-analysis: 33 (11/11/11) <sup>c</sup>	15%	Individuals from hospitals and the community with binge eating disorder. Assessed: PHQ, follow up interview with psychiatric nurse using EDE to confirm whether they met DSM-IV binge eating disorder criteria.	Hospitals, eating disorder community centres, online	82.9% Condition breakdown not given	45 (15) Condition breakdown not given	Canada	Caucasian: 76.6%
Kelly et al., 2017	22 (11/11)	22 (11/11) <sup>a</sup> Observed: 16 (not stated) Meta-analysis: 17 (9/8) <sup>c</sup>	27%	Individuals with a diagnosis of an eating disorder receiving outpatient individual therapy at an eating disorders treatment Centre. Assessed: Semi-structured interview with psychiatric nurse to confirm met DSM-IV criteria for eating disorder.	Outpatient eating disorder treatment centre	100%	Int: 36.7 (12.6) Con: 27.1 (10.1)	USA	Caucasian: 100%
Kelman et al., 2018	123 (61/62)	123 (61/62)	0%	Perinatal women or women intending to become pregnant. Assessed: N/A	Amazon mechanical Turk, professional networks.	100%	61.9% of total sample aged 26-34. Means not given.	USA	European American/White: 53.6% Asian: 28.6%, African American: 8.3%, Latino: 6% Other: 2.4% Southeast Asian: 1.2%

TABLE 2 (Continued)

Study	N recruited (Int/Con)	N analysed post intervention (Int/Con)	Attrition rate (% dropout)	Participant population	Recruitment sources	Gender (% female)	Mean age (SD)	Country	Ethnicity
Kirby & Baldwin, 2018	61 (31/30)	61 (31/30)	0%	Parents with at least one child aged 2-12. Assessed: N/A.	University research participation scheme, advertisement, word of mouth	82%. Condition breakdown not given	38.4 (6.1). Condition breakdown not given	Australia	White Australians: 89%, European: 6%, Asian: 5%
Kirby & Laczko, 2017	97 (49/46)	97 (49/46)	0%	University students living at home. Assessed: N/A.	University courses	Int: 78% Con: 86%	Int: 18.9 (1.6) Con: 18.6 (1.3)	Australia	White: 62%, Asian: 31.6%, Other: 6.3%
Krieger et al., 2019	122 (60/62)	121 (59/62) <sup>a</sup> Observed: 108 (47/61)	11%	General community sample with high level of self-criticism. Assessed: Interviews and FSCRS: IS $\geq$ 20.	Adverts in newspaper articles about e-mental health and on self-help forums	Int: 85% Con: 71%	Int: 38.0 (12.0) Con: 37.4 (11.0)	Switzerland	Not stated
Matos et al., 2017	117 (not reported)	93 (56/37)	21%	General community sample. Assessed: Brief interview to exclude major psychiatric problems and illnesses.	University mailing list	90% Condition breakdown not given	23.3 (4.2) Condition breakdown not given	Portugal	Not stated
Mosewich et al., 2013	60 (31/29)	51 (29/22)	15%	Varsity women athletes with levels of self-criticism that were 'less than constructive', but not interfering with daily functioning. Assessed: Questions at recruitment.	Sports team visits, emails, announcements through coaches, poster displays from universities in Western Canada	100%	Int: 20.3 (2.3) Con: 20.3 (1.1)	Canada	Caucasian: 92%, Black: 4%, Aboriginal: 2%, Chinese: 2%, West Asian: 2%
Palmeira et al., 2017	73 (36/37)	59 (27/32)	19%	Women with overweight and obesity, without binge eating, enrolled in primary care and hospital nutritionist treatment for weight loss. Assessed: Body mass index: body $\geq$ 25, binge eating disorder and severe psychiatric problems excluded through EDE interview, SCID-I and SCID-II interviews.	Discussion with researcher at their medical care unit	100%	Int: 42.0 (8.8) Con: 42.7 (8.4)	Portugal	Not stated
Shahar et al., 2015	38 (19/19)	32 (14/18)	16%	General community sample with high levels of self-criticism. Assessed: DAS-SCP $\geq$ 30, brief phone interview to exclude bipolar/psychotic disorders and self-harm.	Internet adverts, flyers	Int: 74% Con: 47%	Int: 28.7 (10.4) Con: 32.6 (47.4)	Israel	Not stated

(Continues)

TABLE 2 (Continued)

Study	N recruited (Int/Con)	N analysed post intervention (Int/Con)	Attrition rate (% dropout)	Participant population	Recruitment sources	Gender (% female)	Mean age (SD)	Country	Ethnicity
Sommers-Spijkerman, Trompeter, Schreurs, & Bohlmeijer, 2018	243 (121/122)	242 (120/122) <sup>a</sup> Observed: 215 (107, 108)	12%	General community sample with low-moderate well-being. Assessed: MHC-SF: Flourishing well-being excluded, HADS >11 excluded.	Newspapers,	Int: 80% Con: 90%	Int: 52.8 (9.8) Con: 52.9 (10.2)	Netherlands	Dutch: 100%
Stevenson et al., 2019	119(60/59)	119(60/59) <sup>b</sup> Observed: 98 (44, 54) Meta-analysis: 99 (45/54) <sup>c</sup>	18%	General community sample with high levels of social anxiety. Assessed: $\geq 19$ , SPIN.	University online recruitment system, online forums	Int: 75% Con: 78%	Int: 30.9 (12.4) Con: 21.1 (10.6)	Australia	White: 69.7%, Asian: 21.8%, other: 8.5%

Note: Bold font = N used in the meta-analysis. Acronyms from this table in chronological order: Int = Intervention group, Con = Control group, SCS = Self-Compassion Scale (Neff, 2003a), DSM-IV = The Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2013), ICD-10 = International Statistical Classification of Diseases and Related Health Problems (World Health Organization, 2004), PHQ = Personal Health Questionnaire (Spitzer et al., 1999), EDE-Q = Eating Disorder Examination Questionnaire (Fairburn & Beglin, 1994), FSCRS-IS = FSCRS: Inadequate-Self subscale (Gilbert et al., 2004), SCID-I = Structured Clinical Interview for DSM-IV Axis I Disorders (First et al., 1997), SCID-II = Structured Clinical Interview for DSM-IV Axis II Disorders (Lamers et al., 2011), HADS = Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983), SPIN = Social Phobia Inventory (Connor et al., 2000), (de Graaf et al., 2009), MHC-SF = Mental Health Continuum- Short Form (Lamers et al., 2011), HADS = Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983), SPIN = Social Phobia Inventory (Connor et al., 2000).

<sup>a</sup>Intention to treat analysis.

<sup>b</sup>Insufficient data reported so unable to include Halamová, Kanovsky, Varšová, et al. (2018) in meta-analysis.

<sup>c</sup>N reported in paper different to N author gave for data used in meta-analysis.



majority groups. All studies sampled working age adults, with mean sample ages ranging from 18 to 45 years.

Table 3 illustrates the range of intervention durations (one session–14 weeks of support) and delivery methods (such as group, individual, direct and remote self-help) and shows a large proportion of interventions to be informed by Gilbert's (2010) CFT model. Eight of the 19 studies included in the meta-analysis included a follow-up. However, five of these follow-ups no longer had control comparison groups. This is because the follow-ups occurred after the waitlist control groups had received a delayed intervention. Therefore, the data were insufficient to enable this meta-analysis to look at whether intervention effects were maintained at follow-up relative to comparison controls.

Table 3 also shows the range of self-criticism measures used across the studies, the majority showing acceptable internal consistency. The most commonly employed measures were the self-criticism subscales of the Forms of Self-Criticizing/Attacking and Self-Reassuring Scale (FSCRS) (Gilbert et al., 2004), used in 11 of the studies. The FSCRS has demonstrated high internal consistency, adequate test–retest reliability and concurrent validity (Halamová, Kanovský, Varšová, et al., 2018). It assesses forms of self-attacking in response to difficult situations using two subscales: hated-self (FSCRS-HS) and inadequate-self (FSCRS-IS) (Gilbert et al., 2004). The second most commonly employed measure was the negative domains of the SCS or SCS-SF (Neff, 2003a), used in nine of the studies. The SCS and SCS-SF are validated measures, quantifying three negative components of self-compassion (SCS-Neg): self-judgement (SJ), isolation (I) and overidentification (OI) (Neff, 2003a; Raes et al., 2011). The SCS-SJ, SCS-I, SCS-OI and SCS-SF-Neg all have acceptable levels of correlation with FSCRS-HS ( $r = 0.48–0.69$ ) and FSCRS-IS ( $r = 0.42–0.67$ ), illustrating that they quantify self-criticism in a comparable way (Castilho et al., 2015; Sommers-Spijkerman, Trompeter, ten Klooster, et al., 2018).

The remaining six measures were only used once across the studies. Similar to the FSCRS, the Levels of Self-Criticism Scale (LOSC) is a validated measure, measuring self-criticism as a multidimensional construct (Thompson & Zuroff, 2004). Correlational analyses have demonstrated a significant relationship between the LOSC and FSCRS-HS ( $r = 0.55–0.57$ ) and FSCRS-IS ( $r = 0.63–0.77$ ) (Gilbert et al., 2004). The DAS-SCP is a subscale measuring self-criticism in relation to self-critical performance evaluation with good psychometric properties (de Graaf et al., 2009; Weissman & Beck, 1978). Although the subscale has not been formally compared with the most frequently used measures of self-criticism, direct communication with the authors revealed that the DAS-SCP correlated with the FSCRS-IS ( $r = 0.51$ ) and FSCRS-HS ( $r = 0.43$ ) sufficiently in Shahar et al.'s (2015) study to enable results to be compared. The HINT measures the tendency to repetitively criticize oneself as a mental habit (Verplanken et al., 2007). It correlates specifically with the occurrence of negative self-thoughts, not negative thoughts in general (Verplanken et al., 2007). This supports its use as a measure of self-criticism rather than negative thinking more generally. Furthermore, the SSCM is a monitoring tool adapted by Mosewich et al. (2013) where participants

rate frequency, power, intrusiveness, length and feelings associated with a self-critical thought in relation to a recent sporting set-back or failure. Although there is no extensive psychometric support for this measure, when mean levels of self-critical thought were high, self-soothing was low (Gilbert & Procter, 2006).

Finally, despite limited correlational analysis, the content of the S-Con and FFMQ-SJ appear comparable with the measures already discussed. S-Con measures self-condemnation in relation to an offence through four items: bad person, I deserve to suffer for this, angry at myself and hateful towards myself, which is consistent with the conceptualization of the FSCRS-HS (Fisher & Exline, 2006). In addition, the FFMQ-SJ is a validated subscale measuring one's ability to criticize and judge the self harshly (Baer et al., 2006). Overall, the information and evidence available indicates that although the concept of self-criticism can be measured in many different ways, measures are considered similar enough to be compared.

### 2.3 | Study quality

Study quality was not part of the inclusion/exclusion criteria. Instead, issues of quality within the research were considered in relation to how risk of bias may impact the interpretation of results and moderate outcomes. Therefore, the quality of the studies was assessed using a modified version of the Cochrane Collaboration's tool for assessing risk of bias. This assessment tool was chosen because it provides a standard framework for assessing whether there is low, high or unclear risk of bias within RCTs (Higgins et al., 2011). The following domains were assessed: (1) selection bias, (2) performance bias, (3) detection bias, (4) attrition bias, (5) reporting bias, (6) statistical bias and (7) measurement bias. Statistical bias and measurement bias were taken from the National Institute for Health and Care Excellence (NICE) (2012) quality appraisal checklist for quantitative studies and added to the tool to strengthen the global assessment. A qualitative appraisal of these factors in each of the studies is discussed in Section 3. Ten percent of the included papers were independently co-rated, which showed perfect agreement on all items in the coding scheme.

### 2.4 | Information extraction for meta-analysis

For the meta-analysis, self-criticism outcome data were extracted from each paper for intervention and control groups (means, standard deviations, paired *t* values for change, sample size, pre-post correlations). When intention-to-treat statistics were given by the authors, the full sample at randomization was used. When per-protocol results were given by the authors, the sample size for completers was inputted. Therefore, study weightings in the meta-analysis were proportional to the amount of data inputted.

For the moderator analysis, a coding scheme was developed to extract the necessary data. The scheme coded self-criticism type in two forms: hated-self, measured using the FSCRS-HS; and

TABLE 3 Study characteristics

Study	Intervention content	Intervention duration	Intervention delivery	Control	Time points	Self-criticism measures	Cronbach's alpha for questionnaire	Measure translations and/or adaptations
Arimitsu (2016)	Group therapy enhancing self-compassion programme informed by Gilbert's (2010) CFT. <b>Exercises:</b> LKM, mindfulness, compassionate imagery, letter writing, three-chair work etc.	Length: 1.5 h Amount: 7 weekly sessions	Group therapy	WL	Pre Post 3 months FU	SCS-SJ	SCS- Total = 0.88	Japanese (Arimitsu, 2014)
Ascone et al. (2017)	Compassion focused imagery informed by Gilbert's (2010) CFT, read individually to participants by the researchers. <b>Exercises:</b> Compassionate imagery.	Length: 10 min Amount: 1	Individual guided imagery	Active: Individual controlled imagery	Pre Post	FSCRS-IS + HS	0.80	German Subscales shortened, 5 point scale increased to 11
Cornish & Wade, 2015	Individual counselling intervention informed by emotion focused therapy (Greenberg, 2002) promoting forgiveness of others and self. <b>Exercises:</b> Acceptance work, two-chair work, reconnecting with values, self-forgiveness, letter writing etc.	Length: 50 min Amount: 8 weekly sessions	Individual therapy	WL	Pre Post 2 months FU <sup>a</sup>	S-Con	0.83	N/A
Duarte et al., 2017	Group psychoeducation presentation informed by Kabat-Zinn's (1990) mindfulness teaching and Gilbert's (2010) CFT. <b>Exercises:</b> Compassionate imagery, mindfulness, soothing rhythm breathing etc. <b>Post introduction self-practice resources:</b> Written and audio files.	Length: 2.5 h Amount: 1 + 4 weeks self-practice	Group introduction + remote self-practice	WL	Pre Post 1 month FU <sup>a</sup>	FSCRS-IS FSCRS- HS	FSCRS- Total = 0.93	Portuguese (Castilho et al., 2015)
Dundas et al., 2017	Group therapy informed by Neff and Germer's (2013) mindful self-compassion course, Gilbert and Procter's (2006) CMT and Santorelli and Kabat-Zinn's (2009) mindfulness-based-stress reduction. <b>Exercises:</b> Mindfulness, activate soothing system, LKM. <b>Between session self-practice resources:</b> Audio files.	Length: 1.5 h Amount: 3 sessions over 2 weeks + self-practice between	Group therapy	WL	Baseline (3 weeks prior to pre) Post 6 months FU <sup>a</sup>	HINT FFMQ- NJ	0.94 0.84	Norwegian
Feliu-Soler et al., 2017	Group therapy informed by DBT, Gilbert's (2010) and Neff and Germer's (2013) compassion models.	Length: Not stated Amount: 3 weekly sessions + self-practice between	10 weeks of mindfulness group training + group therapy	10 weeks of mindfulness group training + Active: Mindfulness	Pre Post	FSCRS-IS FSCRS-HS	0.90 0.86	Unclear, assumed Spanish

TABLE 3 (Continued)

Study	Intervention content	Intervention duration	Intervention delivery	Control	Time points	Self-criticism measures	Cronbach's alpha for questionnaire	Measure translations and/or adaptations
	<p>Exercises: LKM, compassion meditations, DBT techniques, compassionate letters etc.</p> <p><b>Between session self-practice resources:</b> Audio files.</p>			group continuation				
Halamová, Kanovský, Varšová, et al., 2018 <sup>b</sup>	<p>Emailed self-help emotion focused training for self-compassion and self-protection. Resources informed by Emotion focused therapy, CMT (Gilbert, 2010) and Mindful Self-Compassion Programme (Neff &amp; Germer, 2013).</p> <p><b>Exercises:</b> Self-compassionate letter, letting go of painful memory letters, self-protective anger, self-compassionate mirror, compassionate friend etc.</p> <p><b>Self-practice resources:</b> Emailed text instructions, online documents, reflection questions</p>	<p>Length: 30 min</p> <p>Amount: Daily exercises for 2 weeks</p>	Remote self-practice	WL	PrePost 2 months FU	FSCRS-IS FSCRS-HS SCS-Neg	Not stated Not stated Not stated	Slovak (Halamová et al., 2017)
Johnson et al., 2018	<p>Compassion meditation group informed by Ozawa-de Silva and Negi's (2013) cognitively-based compassion training protocol which stems from Tibetan Buddhism.</p> <p><b>Exercises:</b> Mindful attention, cultivating self-compassion, cultivating compassion for others, appreciation, empathy etc.</p> <p><b>Between sessions self-practice resources:</b> Practice meditation assignments</p>	<p>Length: 1.5 h</p> <p>Amount: 6 weekly sessions + self-practice between</p>	Group therapy	Active: 6 session support group	Pre Post	LOSC	0.82	N/A
Kelly & Carter, 2015	<p>Self-help informed by Binge Eating Disorder CBT self-help materials and compassion-focused therapy (Gilbert, 2010; Goss, 2011).</p> <p><b>Exercises:</b> Compassionate self and compassionate other imagery, letter writing, self-talk etc.</p> <p><b>Self-practice resources:</b> Link to an online form to guide their imagery and letter writing.</p>	<p>Length: Not stated</p> <p>Amount: 1 self-help induction PowerPoint + daily exercises for 3 weeks.</p>	Remote self-practice	Active: behavioural strategies WL	Pre Mid (Week1) Mid (Week2) Post (Week3)	SCS-Neg	SCS-Total = 0.94	N/A

(Continues)

TABLE 3 (Continued)

Study	Intervention content	Intervention duration	Intervention delivery	Control	Time points	Self-criticism measures	Cronbach's alpha for questionnaire	Measure translations and/or adaptations
Kelly et al., 2017	Group focusing on shame and self-criticism, based on Kelly and Leybman (2012) which was informed by Gilbert's (2010) CFT. Exercises: Compassionate other and self-imagery, self-compassionate behaviours, receiving compassion from others, letter writing, thought records etc.	Length: 1.5 h Amount: 12 weekly sessions	Group therapy + treatment as usual (individual CBT/DBT and psychiatry/nutrition)	WL: Out-patient treatment as usual	Pre Mid (Week 4) Mid (Week 8) Post (Week 12)	SCS-Neg	0.89	N/A
Kelman et al., 2018	Didactic internet course based on CFT informed by Gilbert et al. (2010). Exercises: Cultivating compassionate self, compassionate self, details unclear.	Length: 45 min Amount: 1 induction course	Remote self-practice	Active: Internet based CBT	Pre Post	FSCRS-IS FSCRS-HS	Not stated Not stated	N/A
Kirby & Baldwin, 2018	LKM informed by Hofmann et al. (2011) and Germer (2009). Exercises: LKM, mindfulness. Self-practice resources: Audio files.	Length: 15 min Amount: 1	Remote self-practice	Active: Focused imagery	Post	SCS-SJ SCS-SJ-P	Each SCS subscale = 0.85–0.96	Adapted SCS-SJ subscale to focus on parenting and used alongside a parent-child vignette (SCS-SJ-P)
Kirby & Laczko, 2017	LKM informed by Hofmann et al. (2011) and Germer (2009). Exercise: LKM, mindfulness. Self-practice resources: Audio files.	Length: 15 min Amount: 1	Remote self-practice	Active: Focused imagery	Post	SCS-SJ SCS-SJ-C	Each SCS subscale = 0.77–0.83	Adapted SCS-SJ subscale to focus on young adult-parent conflict and used alongside a conflict vignette (SCS-SJ-C)
Krieger et al., 2019	Internet-based self-help based on Mindfulness Based Compassionate Living (Van de Brink & Koster, 2015). Exercises: Mindfulness, LKM, compassionate companion, compassionate breathing, compassionate letter writing, etc. Self-practice resources: Written texts, audio files, diaries.	Length: 50–60 min Amount: 1 module per week for 7 weeks	Remote self-practice + treatment as usual	WL: Treatment as usual	Pre Post 6 months FU <sup>a</sup>	FSCRS-IS FSCRS-HS SCS-Neg	0.58 0.65 SCS-Total = 0.87	German
Matos et al., 2017	Group CMT introduction session informed by Gilbert et al. (2010) followed by remote self-practice. Exercises: Soothing rhythm breathing, friendly facial expression, mindfulness, self-compassion exercise, imagery of compassionate other, self-compassionate exercises etc. Self-practice resources: Written manual, audio files	Length: 2 h Amount: 1 induction session + 2 weeks self-practice	Group introduction + remote self-practice	WL	Pre Post	SCS-SJ FSCRS-HS + IS	SCS subscales = 0.75–0.81 IS = 0.9 HS = 0.86	Unclear, assumed Portuguese

TABLE 3 (Continued)

Study	Intervention content	Intervention duration	Intervention delivery	Control	Time points	Self-criticism measures	Cronbach's alpha for questionnaire	Measure translations and/or adaptations
Mosewich et al., 2013	Group psychoeducational induction presentation followed by writing tasks informed by Leary et al. (2007). <b>Exercises:</b> Compassion writing tasks based on common humanity, self-kindness and mindfulness etc. <b>Self-practice resources:</b> Written materials.	Length: 10 min Amount: 1 induction session + 5 self-practice exercises over 7 days	Group introduction + remote self-practice	Active: Attention control	Pre Post 1 month FU	SSCM	Not stated	Modified to focus on a recent sporting event
Palmeira et al., 2017	Kg-Free group therapy intervention was informed by Gilbert's (2010) CFT, ACT and mindfulness techniques to target weight self-stigma and unhealthy eating behaviours. <b>Exercises:</b> Mindfulness, action, acceptance work, cognitive diffusion, soothing rhythm breathing, urge surfing, safe place, compassion letter writing, compassionate-self exercise etc. <b>Between sessions self-practice resources:</b> Audio files, written manual	Length: 2.5 h Amount: 10 weekly sessions + 2 biweekly booster sessions	Group therapy	WL: Treatment as usual	Pre Post	FSCRS-IS FSCRS-HS	0.79 0.64	Unclear, assumed Portuguese
Shahar et al., 2015	LKM group informed by Salzberg (1995) and lead by Vipassana meditation teacher. <b>Exercises:</b> Compassion towards self, compassion towards others, gratitude etc. <b>Between sessions self-practice resources:</b> CD audio exercises	Length: 1.5 h Amount: 7 weekly sessions	Group therapy	WL	Pre Post 3 months FU <sup>a</sup>	FSCRS-IS FSCRS-HS DAS-SCP	0.88 0.64 0.89	Unclear, assumed Hebrew
Sommers-Spikerman, Trompeter, Schreurs, & Bohlmeijer, 2018	Self-help book informed by Gilbert's (2010) CFT. <b>Exercises:</b> Mindful breathing, self-criticism thought diary, ideal compassionate self, compassionate letter to other etc. <b>Self-practice resources:</b> Self-help book, audio exercises.	Length: Not stated Amount: 7 weekly lessons over 9 weeks	Remote self-practice	WL	Pre Post 3 months FU 9 months FU	FSCRS-IS FSCRS-HS SCS-SF-Neg	Not stated Not stated Not stated	Unclear, assumed Dutch

(Continues)

TABLE 3 (Continued)

Study	Intervention content	Intervention duration	Intervention delivery	Control	Time points	Self-criticism measures	Cronbach's alpha for questionnaire	Measure translations and/or adaptations
Stevenson et al., 2019	Online CFT informed by Gilbert's (2010) model. Exercises: Compassionate letters to self-regarding recent social anxiety scenario	Length: 5–15 min Amount: Daily tasks for 2 weeks	Remote self-practice	Active: Cognitive restructuring	Pre Mid Post 1weekFU 5weekFU	FSCRS-HS FSCRS-IS SCS-Neg	FSCRS subscales = 0.94–0.97 SCS-Neg subscales = 0.9–0.94	N/A

Note: General acronyms from Table 3 in chronological order: CFT = compassion-focused therapy, LKM = Loving Kindness Meditation, WL = Waitlist, FU = Follow-up, CMT = Compassionate Mind Training, CBT = Cognitive Behavioural Therapy, DBT = Dialectical Behavioural Therapy, ACT = Acceptance and Commitment Therapy, Questionnaire acronyms from Table 3 in alphabetical order: DAS-SCP = Dysfunctional Attitude Scale: Self-Critical Perfectionism subscale (de Graaf et al., 2009), FFMQ-NJ = Five Facet Mindfulness Questionnaire: Non-judgment subscale (Baer et al., 2006), FSCRS = Forms of Self-Criticism/Attacking and Self-Reassurance Scale (Gilbert et al., 2004), FSCRS-IS = FSCRS: Inadequate-Self subscale (Gilbert et al., 2004), FSCRS-IS + HS = FSCRS: Inadequate-Self and Hated Self subscales combined (Gilbert et al., 2004), FSCRS-HS = FSCRS: Hated-Self subscale (Gilbert et al., 2004), HINT = Habit Index of Negative Thinking (Verplanken et al., 2007), SSCM = State Self-Criticism Measure (Gilbert & Procter, 2006; Mosewich et al., 2013), S-Con = Self-Compassion Exhale, 2006), SCS-Total = Self-Compassion Scale: All 6 items (Neff, 2003a), SCS-Neg = Self-Compassion Scale: Negative items (Self-Judgement, Isolation, Overidentification) (Neff, 2003a), SCS-SF-Neg = Self-Compassion Scale Short Form: Negative items (Neff, 2003a), SCS-SJ = Self-Compassion Scale: Self-judgment focusing on parenting (Kirby & Baldwin, 2018), LOSC = Levels of Self-Criticism Scale (Thompson & Zuroff, 2004), SJP-P = Self-Compassion Scale: Self-judgment focusing on parenting (Kirby & Baldwin, 2018), LOSC = Levels of Self-Criticism Scale (Thompson & Zuroff, 2004).  
<sup>a</sup>Control received delayed intervention so before FU so no control group at FU.  
<sup>b</sup>Insufficient data reported so unable to include Halamová, Kanovský, Varšová, et al. (2018) in meta-analysis.

inadequate-self, measured using the FSCRS-IS (Gilbert et al., 2004). There were insufficient data to include any other conceptualizations of self-criticism types measured amongst the papers in the moderator analysis. This is because all other measures or subscales of self-criticism were reported in four studies or fewer. The scheme also included intervention duration in days and sample type (clinical or nonclinical). Samples were identified as 'clinical' if they consisted of individuals with diagnosed mental health difficulties and/or who were patients of mental health services. Samples were labelled 'nonclinical' if they consisted of individuals without mental health diagnoses in university, healthcare or community settings.

Finally, the scheme coded for control comparison group (passive, including waitlist or treatment as usual, or active), intervention delivery (remote, such as online self-help, or direct face to face intervention with a therapist), intervention setting (individual or group therapy) and risk of bias (high/uncertain risk or low risk). A study was classified as 'high/uncertain risk of bias' if it was rated as 'high' or 'uncertain' on 50% or more of the domains on the modified risk of bias Cochrane Collaboration tool. Conversely, a study was rated as 'low risk of bias' if it was rated 'low' on more than 50% of the assessment domains. 'High' and 'uncertain' were combined because it was thought a domain was at high risk of bias if it had not openly reported the information of interest. Table 4, which is available in e-version, shows how each study was coded.

## 2.5 | Effect size calculation

Comprehensive Meta-Analysis (CMA) Version 3 software was used to calculate effect sizes (Borenstein et al., 2011). A random effects model was used because there was variability in the study designs and locations (Borenstein et al., 2011). Effects sizes were coded so that a positive effect size indicated that self-compassion-related interventions were more effective at reducing self-criticism than control conditions. Conversely, a negative effect size indicated that control conditions were more effective than self-compassion-related interventions at reducing self-criticism.

The majority of studies reported participant numbers, means and standard deviations for each group which were inputted into CMA. In order to gain the necessary pre-post within group change  $t$  values and pre-post correlation  $r$  values, it was necessary to contact the majority of the authors. Despite this, pre-post correlation values were not available for four studies (Arimitsu, 2016; Duarte et al., 2017; Feliu-Soler et al., 2017; Johnson et al., 2018). Sensitivity analysis inputting pre-post correlation  $r$  values ranging from 0.1 to 0.9 for the four studies missing correlations showed that the overall effects of the meta-analysis remained unchanged. The significance of the categorical moderators was also not affected by the possible range of the missing pre-post correlation values. As all the studies were of a similar RCT design, it was felt a mean pre-post correlation value, calculated from the 15 studies supplying  $r$  values ( $M = 0.65$ ), provided a realistic estimate for the four missing pre-post correlations (Higgins et al., 2020). Therefore, 0.65 was inputted as a substitute

**TABLE 4** Moderator coding for meta-analysis

Study	Forms of self-criticism	Intervention length (days)	Sample type	Control comparison group	Intervention delivery	Intervention setting	Risk of bias
Arimitsu, 2016	Not included <sup>a</sup>	49	Nonclinical	Passive	Direct	Group	High/unclear
Ascone et al., 2017	Not included <sup>a</sup>	1	Clinical	Active	Direct	Individual	High/unclear
Cornish & Wade, 2015	Not included <sup>a</sup>	56	Nonclinical	Passive	Direct	Individual	High/unclear
Duarte et al., 2017	Inadequate-self Hated-self	28	Clinical	Passive	Not included <sup>b</sup>	Not included <sup>b</sup>	High/unclear
Dundas et al., 2017	Not included <sup>a</sup>	14	Nonclinical	Passive	Direct	Group	Low
Feliu-Soler et al., 2017	Inadequate-self Hated-self	21	Clinical	Active	Direct	Group	High/unclear
Johnson et al., 2018	Not included <sup>a</sup>	42	Clinical	Active	Direct	Group	High/unclear
Kelly & Carter, 2015	Not included <sup>a</sup>	21	Clinical	Active	Remote	Individual	High/unclear
Kelly & Carter, 2015	Not included <sup>a</sup>	21	Clinical	Passive	Remote	Individual	High/unclear
Kelly et al., 2017	Not included <sup>a</sup>	84	Clinical	Passive	Direct	Group	High/unclear
Kelman et al., 2018	Inadequate-self Hated-self	1	Nonclinical	Active	Remote	Individual	Low
Kirby & Baldwin, 2018	Not included <sup>a</sup>	1	Nonclinical	Active	Remote	Individual	Low
Kirby & Laczko, 2017	Not included <sup>a</sup>	1	Nonclinical	Active	Remote	Individual	Low
Krieger et al., 2019	Inadequate-self Hated-self	49	Nonclinical	Passive	Remote	Individual	Low
Matos et al., 2017	Not included <sup>a</sup>	14	Nonclinical	Passive	Not included <sup>b</sup>	Not included <sup>b</sup>	High/unclear
Mosewich et al., 2013	Not included <sup>a</sup>	7	Nonclinical	Active	Not included <sup>b</sup>	Not included <sup>b</sup>	Low
Palmeira et al., 2017	Inadequate-self Hated-self	98	Nonclinical	Passive	Direct	Group	Low
Shahar et al., 2015	Inadequate-self Hated-self	49	Nonclinical	Passive	Direct	Group	Low
Sommers-Spijkerman, Trompetter, Schreurs, & Bohlmeijer, 2018	Inadequate-self Hated-self	63	Nonclinical	Passive	Direct	Group	Low
Stevenson et al., 2019	Inadequate-self Hated-self	14	Nonclinical	Active	Direct	Group	Low

<sup>a</sup>Study did not report FSCRS-IS or FSCRS-HS scores.

<sup>b</sup>Intervention unable to be coded as direct/remote or individual/group due to involving both a group introduction and a remote individual self-practice.

$r$  for the studies missing pre-post correlations. However, sensitivity analysis showed that the significance of the continuous moderator, intervention duration, was sensitive to the possible values of the four missing pre-post correlations. When inputted  $r$  values for Arimitsu (2016), Duarte et al. (2017), Feliu-Soler et al. (2017) and Johnson et al. (2018) were  $>0.87$ , the significance of the intervention duration moderator was changed. This is taken into consideration in the interpretation of the findings.

Hedges'  $g$  values calculated by CMA are reported, a variation of Cohen's  $d$  that corrects for bias due to small sample sizes (Hedges & Olkin, 1985). The significance of effect sizes was determined based on the  $p$  values of the resultant  $Z$  scores. When a study reported two measures of self-criticism, a CMA-calculated

mean of the outcomes was used to prevent greater weight being attributed to studies with multiple outcomes using the same sample. However, when studies had more than two measures of self-criticism (Krieger et al., 2019; Shahar et al., 2015; Sommers-Spijkerman, Trompetter, Schreurs, & Bohlmeijer, 2018; Stevenson et al., 2019), composite values were computed outside of CMA using the actual correlation  $r$  values between measures. These values were gained through communication with the authors in order to calculate a more precise estimate of variance, in line with Borenstein et al.'s (2011) recommendations and formula. It was not possible to gain the correlation between measures for Stevenson et al.'s (2019) paper. Therefore, mean correlation between measures  $r$  values were calculated from Krieger et al. (2019),

**TABLE 5** Risk of bias assessment

Author and year	Selection	Bias	Performance bias	Detection bias	Attrition bias	Reporting bias	Statistical bias	Measurement bias
	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Sufficient statistical power	Psychometric properties of measures
Arimitsu, 2016	Unclear	Low	High	Unclear	Low	High	Low	Low
Ascone et al., 2017	Low	Low	High	High	Unclear	Low	High	Low
Cornish & Wade, 2015	Low	Low	High	Unclear	High	Low	High	Low
Duarte et al., 2017	Unclear	Low	High	Unclear	High	Low	High	Low
Dundas et al., 2017	Low	Low	High	Unclear	Low	Low	Low	Low
Feliu-Soler et al., 2017	Low	Unclear	High	Unclear	Unclear	Low	High	Low
Halamová, Kanovský, Varšová, et al., 2018 <sup>a</sup>	Low	Low	High	Unclear	High	Low	Unclear	Low
Johnson et al., 2018	Unclear	Low	High	Unclear	Unclear	Low	Unclear	Low
Kelly & Carter, 2015	Low	Low	High	Unclear	Low	High	Unclear	Low
Kelly et al., 2017	Low	Low	High	Unclear	High	Low	Unclear	Low
Kelman et al., 2018	Low	Low	High	Unclear	Low	Low	Low	Low
Kirby & Baldwin, 2018	Low	Low	High	Low	Low	High	Low	Low & unclear
Kirby & Laczko, 2017	Low	Low	High & low	Low	Low	High	High	Low & unclear
Krieger et al., 2019	Low	Low	High	Unclear	Low	Low	Low	High & unclear
Matos et al., 2017	Unclear	High	High	Unclear	High	Low	Low	Low
Mosewich et al., 2013	Low	Unclear	High	Low	Low	Low	Low	Unclear
Palmeira et al., 2017	Low	Low	High	Low	Low	Low	Low	Low & high
Shahar et al., 2015	Low	Low	High	Unclear	Low	Low	Low	Low & high
Sommers-Spijkerman, Trompetter, Schreurs, & Bohlmeijer, 2018	Low	Low	High	Unclear	Low	Low	Low	Unclear
Stevenson et al., 2019	Low	Unclear	High	Unclear	Low	Low	Low	Low

<sup>a</sup>Insufficient data reported so unable to include Halamová, Kanovský, Varšová, et al. (2018) in meta-analysis.



Shahar et al. (2015) and Sommers-Spijkerman, Trompetter, Schreurs, and Bohlmeijer's (2018) studies when the same measures were used and inputted as substitutes.

One study (Kelly & Carter, 2015) compared a self-compassion intervention to two groups: an active control and a waitlist control. In line with Borenstein et al.'s (2011) recommendations, the number of participants in the compassion intervention was divided by the number of comparison conditions to prevent the study weighting the analysis disproportionately to its sample size. Kelly and Carter's (2015) paper was therefore counted and referred to as if it were two separate samples in the meta-analysis.

### 3 | RESULTS

#### 3.1 | Quality appraisal of studies

The risk of bias findings are summarized in Table 5, including Halamová, Kanovský, Varšová, et al. (2018) study, which was not able to be included in the meta-analysis. All 20 of the studies were unable to blind participants to the intervention being received, a common risk of bias in psychological research (Kirby et al., 2017; Wilson et al., 2019). After this, attrition bias and statistical bias were the next highest risk domains, with eight and nine studies respectively rated as 'high' or 'unclear' risk. However, generally the attrition rates were in the expected range for RCTs. Across all studies, the domain most poorly reported was detection bias, with 15 out of the 20 studies being rated as 'unclear' for blinding of outcome assessment. However, selection bias, reporting bias and measurement bias were the lowest risk of bias domains across the studies with 16 or 17 studies being classified as 'low' risk.

Palmeira et al.'s (2017) study was highlighted as the strongest study overall because they had the highest proportion of low risk rated domains (6.5 out of eight domains). Dundas et al.'s (2017) study and Kelman et al.'s (2018) study were noted to be the second most robust studies with six low risk domains. Conversely, Duarte et al. (2017), Feliu-Soler et al. (2017) and Matos et al. (2017) were the studies most likely to be at high risk of bias, with five high/uncertain risk domains. The study most unclearly reported was Johnson et al. (2018), with 50% of domains being rated as 'unclear'.

#### 3.2 | Meta-analysis overall findings

A total of 20 samples from 19 papers, involving 1350 participants, compared a self-compassion-related intervention to a control condition. Using a random-effects model, the meta-analysis indicated a mean weighted effect size of Hedges'  $g = 0.51$ , 95% CI [0.33–0.69],  $p < 0.001$ . This constitutes a medium effect size and indicates that the participants allocated to self-compassion-related interventions evidence greater reduction in self-criticism than those assigned to control conditions.

#### 3.3 | Publication bias

A funnel plot of precision of Hedges'  $g$  was visually inspected to assess risk of publication bias (see Figure 2). This did not indicate an asymmetrical trend of more studies on the right towards the bottom of the plot, which would suggest publication bias, other than one outlier to the bottom left (Borenstein et al., 2011). The number of missing studies with a mean Hedges'  $g$  of 0 that would make the effect size no longer significant, the fail-safe  $N$ , was 395. It is unlikely that such large numbers of studies with null results exist. Furthermore, Duval and Tweedie's (2000) Trim and Fill technique also suggested the risk of publication bias in the sample to be low. This is because when the studies were trimmed from the left side of the mean using a random effects model, there was no change in the effect size. When the studies were trimmed from the right side, the effect size change was small, Hedges'  $g = 0.53$ , 95% CI [0.35–0.72]. When sensitivity analysis was conducted by removing each study and recalculating the statistics, the effects sizes ranged from 0.54, 95% CI [0.36–0.73] in the study with the biggest sample (Kelman et al., 2018;  $N = 123$ ) to 0.45, 95% CI [0.29–0.61] in the study with the second biggest sample size (Krieger et al., 2019;  $N = 122$ ). Overall, this evidence suggests that the findings were not likely to have been heavily influenced by publication bias.

#### 3.4 | Moderator analysis

Homogeneity analyses for the sample indicated that the effects were heterogeneous  $Q(19) = 49.47$ ,  $p < 0.001$ ,  $I^2 = 61.59\%$ . This value suggests that the variability in results was greater than expected based on sampling error associated with the sample sizes (Rosenthal, 1991). Therefore, moderator analysis was conducted. Each moderator was looked at individually because there was not enough power to conduct a meta-analysis looking at all moderators in one model. It is generally understood that for every covariate included in meta-regression, 10 studies are required (Borenstein et al., 2011). As this study was interested in seven moderators, it would have required a minimum of 70 studies to conduct a meta-regression, which was not feasible.

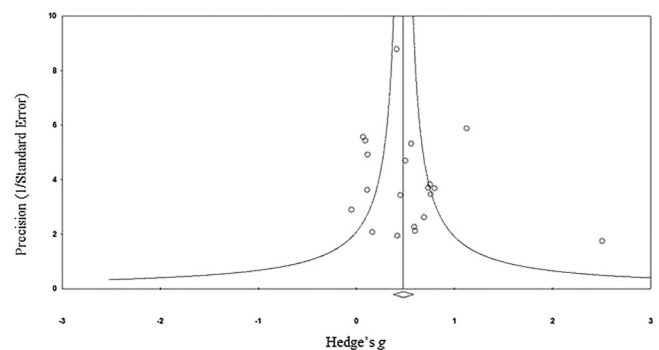


FIGURE 2 Funnel plot of precision of Hedges'  $g$

Despite multiple comparisons inflating the risk of a type one error, it was decided that no correction would be applied to the moderator  $p$  values. This is due to corrections, such as the Bonferroni (Hedges & Olkin, 1985) or Holm (1979) correction, frequently being considered as overly conservative in moderator analysis and inflating type two error (Wang & Ware, 2013). As an alternative to a formal adjustment for multiple comparisons, any significant moderator findings were planned to be discussed in context of the number of tests performed (Borenstein et al., 2011; Wang & Ware, 2013).

Length of intervention was investigated as a continuous moderator using meta-regression. This is because the large variety of intervention durations across studies meant subgroup analysis was considered inappropriate. The meta-regression, using length of intervention as a covariate, found that study findings were significantly related to intervention length,  $Q(1) = 4.15$ ,  $p = 0.042$ ,  $R^2 = 0.15$ , with a medium effect, which explained 15% of the true variance in effects (Cohen, 1988). As the length of an intervention increases by 1 day, the effect size of a study increases by 0.007, illustrating that longer self-compassion-related interventions are more effective at reducing self-criticism than shorter ones,  $\beta = 0.007$  [0.003–0.013],  $p = 0.041$ . Residual heterogeneity was significant  $Q(18) = 41.32$ ,  $p < 0.001$ , indicating that 56.44% ( $I^2$ ) of variability remained unexplained by the model and may potentially be explained by additional moderators.

All remaining moderators were investigated using subgroup analysis due to being categorical in nature. The type of comparison control

group was found to significantly moderate findings,  $Q(1) = 5.53$ ,  $p = 0.019$ , with self-compassion-related interventions showing a larger effect of reducing self-criticism relative to a passive control than an active control. The remaining five moderators were found to have no significant influence on findings (see Table 6).

### 3.5 | Power

Using Hempel et al.'s (2013) guidelines to produce a retrospective power estimate, it is estimated that the four nonsignificant subgroup analyses each had power of around 0.2 to detect a medium moderator effect (number of studies included = 10–20, sample size within trials = 50–100, moderator effect = 0.2, degree of residual heterogeneity  $\tau^2 = 0.1$ , balance of trial level moderator = 25%–50%). The accepted power threshold is 0.8, indicating that the five insignificant moderators were severely underpowered.

## 4 | DISCUSSION

The aim of the review was to estimate the overall effect of self-compassion-related interventions on self-criticism and investigate potential moderating variables. The systematic search identified 19 RCT papers that met the inclusion criteria for the meta-analysis.

**TABLE 6** Moderator effects

Moderator	k	Hedges' g [95% CI]	Z	Heterogeneity
Forms of self-criticism	-	-	-	Between groups: $Q(1) = 0.048$ , $p = 0.826$
Hated-self	8	0.36 [0.07–0.66]	2.47*	$Q(7) = 23.60$ , $p = 0.001^{**}$ , $I^2 = 70.33$
Inadequate-self	8	0.41[0.12–0.71]	2.76**	$Q(7) = 22.77$ , $p = 0.002^{**}$ , $I^2 = 69.26$
Sample type	-	-	-	Between groups: $Q(1) = 1.096$ , $p = 0.295$
Clinical	7	0.38 [0.11–0.65]	2.74**	$Q(6) = 5.51$ , $p = 0.480$ , $I^2 = 0.00$
Nonclinical	13	0.57 [0.34–0.79]	4.87***	$Q(12) = 43.37$ , $p = 0.000^{***}$ , $I^2 = 72.33$
Comparison control group	-	-	-	Between groups: $Q(1) = 5.53$ , $p = 0.019^*$
Active	9	0.30 [0.08–0.52]	2.66**	$Q(8) = 14.00$ , $p = 0.082$ , $I^2 = 42.86$
Passive	11	0.69 [0.45–0.94]	5.56***	$Q(10) = 24.01$ , $p = 0.007^{**}$ , $I^2 = 58.50$
Intervention delivery	-	-	-	Between groups: $Q(1) = 0.02$ , $p = 0.885$
Direct	11	0.50 [0.26–0.74]	4.13***	$Q(10) = 23.87$ , $p = 0.008^{**}$ , $I^2 = 58.11$
Remote	6	0.46 [0.02–0.91]	2.04*	$Q(5) = 24.25$ , $p = 0.000^{***}$ , $I^2 = 79.38$
Intervention setting	-	-	-	Between groups: $Q(2) = 0.34$ , $p = 0.562$
Group	9	0.44 [0.28–0.60]	5.25***	$Q(8) = 9.45$ , $p = 0.306$ , $I^2 = 15.37$
Individual	8	0.58 [0.14–1.02]	2.56*	$Q(7) = 38.45$ , $p = 0.000^{***}$ , $I^2 = 81.80$
Risk of bias	-	-	-	Between groups: $Q(1) = 0.09$ , $p = 0.765$
High/unknown	10	0.55 [0.23–0.88]	3.30**	$Q(9) = 18.80$ , $p = 0.027^*$ , $I^2 = 52.13$
Low	10	0.49 [0.26–0.72]	4.21***	$Q(9) = 30.60$ , $p = 0.000^{***}$ , $I^2 = 70.58$

Note: k = number of studies, Z = Z score, Q = test statistic for heterogeneity,  $I^2$  = measure of degree of heterogeneity. g = 0.2–0.5 = small effect, g > 0.5–0.8 = medium effect, g > 0.8 = large effect (Cohen, 1988).

\* $p < 0.05$ .

\*\* $p < 0.01$ .

\*\*\* $p < 0.001$ .

These papers were from a 6-year period and involved 1350 participants from 20 samples. The meta-analysis indicated that self-compassion-related interventions produce a significant, medium reduction in self-criticism in comparison with control groups. This replicates Ferrari et al.'s (2019) meta-analysis findings in a larger sample of RCTs, suggesting the finding is reliable. Furthermore, greater reductions in self-criticism were seen when self-compassion-related interventions were longer in duration of days and compared with passive controls rather than active. The remaining moderators investigated, self-criticism type, sample type, intervention delivery, intervention setting and risk of bias, were not significant. However, conclusions are qualified by RCTs being of moderate quality and there being a large amount of heterogeneity in effects.

As seven independent moderators were investigated, each at the 0.05  $\alpha$  level, the inflated probability of a type one error in the moderator analysis is recognized. However, as the number of significant moderators was greater than the number expected by chance, the research has more confidence that some of these findings are true (Lagakos, 2006; Wang & Ware, 2013). Greater confidence can be placed in the significance of the control comparison type moderator. This because it had a smaller  $p$  value than intervention length and sensitivity analysis showed only intervention duration moderator was sensitive to the estimated values of the missing pre-post  $r$  for four studies. Thus, it is thought to be unlikely that both the two significant moderators are false positives. However, the risk of one false positive, particularly intervention length, is recognized. The importance of papers routinely reporting pre-post correlation values is highlighted to improve the accuracy of future meta-analyses.

The significance of the control comparison type moderator is supported by previous compassion meta-analyses also reporting control type to significantly moderate a range of psychological outcomes (Ferrari et al., 2019; Wilson et al., 2019). Whereas previous research has only found significant effects when compassion interventions were compared with passive waitlist conditions (Ferrari et al., 2019; Wilson et al., 2019), this study still saw a small positive effect size when self-compassion-related interventions were compared with active control groups. This suggests that self-compassion-related interventions may be able to rival active controls in relation to reducing self-criticism. This is a very promising finding, highlighting the progression of research since Leaviss and Uttley's (2015) early review.

Although less confidence can be placed on the significant moderator of intervention length, it remains important to discuss because the majority of previous meta-analyses have not investigated intervention length as a moderator (Ferrari et al., 2019; Kirby et al., 2017; Wilson et al., 2019). The meta-regression indicated that the longer the intervention, the greater the outcome. Thus, it would be of interest to examine length of intervention in continued research. In order to reduce the effects of confounding variables, it is recommended that future RCTs consider comparing different lengths of the same compassion intervention within their study. This is because intervention length is likely to be influenced by the population and form of self-criticism sampled. For example, the hated-self form of self-criticism is associated with higher levels of psychopathology and is found to be

more persistent and resistant to change than the inadequate-self form (Castilho et al., 2017; Werner et al., 2019). Although the forms of self-criticism and sample type moderators were insignificant, more RCTs in the future could enable a meta-regression to investigate how moderators interact with each other within one model (Hempel et al., 2013).

In relation to the five variables that were not found to significantly moderate effects, it is noted that the retrospective power estimate of the subgroup analyses was low (Hedges & Pigott, 2004; Hempel et al., 2013). To conserve power, only two subgroups within moderators were investigated. However, this is likely to have reduced the sensitivity of the moderator analyses due to higher levels of heterogeneity within subgroups. With more RCTs, a future meta-analysis may have greater power to investigate more specific moderator subgroups. This would particularly benefit sample type understanding as the current categories 'clinical' and 'nonclinical' were extremely broad. In contrary to this review, Ferrari et al.'s (2019) meta-analysis found that the type of participant population and intervention delivery significantly moderated compassion intervention effects on psychological outcomes. Therefore, the failure to reject the majority of the moderators' null hypotheses should be interpreted with caution. In light of this review's limited power and high heterogeneity, the most reasonable conclusion is that there is currently not enough information to adequately judge whether the remaining five moderators have meaningful effects on the reduction of self-criticism in self-compassion-related interventions. With more RCTs, a future meta-regression could improve the power of moderator analyses by reducing the background noise of residual heterogeneity (Hempel et al., 2013).

Furthermore, the review highlights the large range of self-criticism measures being used and the inconsistent reporting of subscales between studies. A more standardized approach to measuring self-criticism would increase the comparability of outcomes across studies for a future review, enabling more precise conclusions to be drawn. Going forward, RCTs are recommended to prioritize the use of the FSCRS (Gilbert et al., 2004) and the SCS (Neff, 2003a) and consistently report their separate subscales. This is because these two measures have been used most frequently in the research to date and are both validated and reliable measures. Through a more consistent approach, it is hoped that future meta-analyses will have greater power to investigate the forms of self-criticism moderator more thoroughly, including not only Gilbert et al.'s (2004) hated-self and inadequate-self forms but also Neff's conceptualization of 'self-judgment' from the SCS (Neff, 2003a), to widen understanding.

#### 4.1 | Risk of bias

Publication bias is a common limitation of meta-analyses. However, the funnel plot analysis, failsafe  $N$  and Trim and Fill technique (Duval & Tweedie, 2000) indicated that the findings were robust and publication bias was unlikely to have heavily influenced the findings. Nevertheless, the validity of the meta-analysis could have been

strengthened by including unpublished literature. Thus, it is recommended that future reviews take this into consideration in order to overcome this limitation.

In agreement with Wilson et al.'s (2019) observations, this review's within study risk of bias assessment highlighted how the general quality of self-compassion-related intervention research has improved since Leaviss and Uttley's (2015) early review. Nevertheless, there were frequent aspects that were not reported clearly within the studies, such as detection bias. This left many 'unclear' ratings in the risk assessment, reducing the accuracy of the risk of bias ratings for each study. This confounding variable is likely to have reduced the validity of the risk of bias moderator analysis. In order to gain a more accurate understanding of how study quality impacts findings, research is required to more vigorously report their methodology so future reviews can more clearly define 'high' and 'low' risk of bias studies.

Furthermore, high performance bias was reported across all studies. Although successfully blinding participants to group allocation is a challenge in psychological research, the risk of performance bias across studies could have been reduced if more active control conditions were used. This is because participants' expectations of two different therapies would be more matched and easier to blind participants to compared with the vast difference in expectations and experiences of a therapy versus a waitlist condition. This provides another argument for future research to prioritize the use of active control conditions in RCTs. However, overall, the 20 studies were of reasonable quality, particularly in relation to showing trends for low selection, reporting and measurement bias

## 4.2 | Further limitations

This meta-analysis distinguished itself from previous reviews by taking a more inclusive, nondiagnostic stance in order to provide a comprehensive overview of current literature and consider moderator effects. However, it is recognized that the limitation of a wide inclusion criteria is that a large amount of variability in samples, measures, interventions and active controls was introduced. This makes identifying the mechanisms of change more complex. However, if RCTs continue to investigate a range of self-compassion-related interventions, it would be interesting for a future meta-analysis to consider compassion intervention type as a moderator, such as CFT compared with interventions based predominately in Neff's conceptualization (2003a, 2003b). This would help to further understand the process of change. Furthermore, the study's broad inclusion criteria enabled interventions such as LKM to be included (Kirby & Baldwin, 2018; Kirby & Laczko, 2017). However, in acknowledgement of Gilbert et al.'s (2019) initial research distinguishing compassion from kindness amongst student and community participants, it is important that future research more closely considers the relationship between LKM and self-compassion to inform future reviews.

It is also important to recognize that the majority of samples were females aged between 18 and 45 years. Research suggests that

women typically report lower levels of self-compassion and higher levels of fear of compassion than men (Xavier et al., 2015; Yarnell et al., 2015, 2019). Therefore, it is likely that the field of self-compassion research is biased by predominately sampling females. A larger effect for self-compassion and lower attrition rates may have been seen if more male samples were included (Gilbert & Procter, 2006). This highlights the need for more gender balanced samples to be recruited in future RCTs. It would also be interesting for studies to investigate the effects of interventions in child, adolescent and older adult samples. This is because research suggests self-compassion is related with age and can have protective effects in these populations (Bluth et al., 2017; Homan, 2016).

Furthermore, the majority of RCTs sampled individuals from White, ethnic majority groups in Western cultures, a common limitation within psychological research (Miranda et al., 2003). This is concerning considering the disproportionate exposure to discrimination and stressful life experiences ethnic minority groups are known to face, and the association between symptoms of depression and ethnic group discrimination (Carr et al., 2014; Finch et al., 2000; Matthews et al., 2013; Noh et al., 1999; Perry et al., 2013). Therefore, individuals from ethnic minority groups may respond differently to interventions due to potentially higher levels of distress and adversity. Additionally, the majority of research in the field is not representative of Eastern cultures, such as Japan where individuals have a tendency to be more self-critical and self-compassion interventions may be less tolerable (Kitayama et al., 1997). In order to improve the generalizability of the findings to impact a wider population of individuals, these underrepresented demographics should be taken into consideration in future research design and recruitment.

Finally, due to the lack of controlled follow-up time points in the research, this review was unable to assess the longitudinal impact of interventions relative to controls. Previous evidence suggests that self-compassion may have buffering effects against development of depression and intervention improvements in depression and self-compassion may be maintained afterwards (Ferrari et al., 2018, 2019). However, more studies with follow-up time points at 6 months or later are required to establish this (Kirby et al., 2017). If RCTs more frequently start using active controls, it will be more feasible for follow-up time points to still have a control comparison, unlike many of the waitlist control designs.

## 5 | CONCLUSION

Overall, this is the largest scale meta-analysis reviewing the effects of self-compassion-related interventions on self-criticism. Findings suggest promising outcomes. However, to reduce performance bias and establish long-term effectiveness relative established therapies, more RCTs using active controls with longitudinal follow-ups are required. This will increase the power of future reviews and may enable a meta-regression investigating moderators with more specific subgroups to be considered. This would enable mechanisms of change to be explored more closely. In order to gain more understanding in how

different forms of self-criticism may moderate effects, RCTs are recommended to routinely report the individual subscales from the FSCRS (Gilbert et al., 2004) and the SCS (Neff, 2003a). Finally, the generalizability of the research field could be advanced by future research sampling a broader range of genders, ages, cultures and ethnicities. In conclusion, the meta-analysis provides evidence that self-compassion-related interventions may successfully be able to target the transdiagnostic, pervasive construct of self-criticism.

## CONFLICT OF INTERESTS

The authors declare that they have no conflicts of interest.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## ORCID

Katherine E. Wakelin  <https://orcid.org/0000-0003-3400-9233>

Gemma Perman  <https://orcid.org/0000-0001-8483-9978>

Laura M. Simonds  <https://orcid.org/0000-0002-8173-4051>

## REFERENCES

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-V)* (5th ed.). Arlington, VA: American Psychiatric Publishing.
- Arimitsu, K. (2014). Development and validation of the Japanese version of the self-compassion scale. *The Japanese Journal of Psychology, 85*, 50–59. <https://doi.org/10.4992/jjpsy.85.50>
- Arimitsu, K. (2016). The effects of a program to enhance self-compassion in Japanese individuals: A randomized controlled pilot study. *The Journal of Positive Psychology, 11*(6), 559–571. <https://doi.org/10.1080/17439760.2016.1152593>
- Ascone, L., Sundag, J., Schlier, B., & Lincoln, T. M. (2017). Feasibility and effects of a brief compassion-focused imagery intervention in psychotic patients with paranoid ideation: A randomized experimental pilot study. *Clinical Psychology & Psychotherapy, 24*(2), 348–358. <https://doi.org/10.1002/cpp.2003C>
- Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J., & Toney, L. R. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment, 13*, 27–45. <https://doi.org/10.1177/1073191105283504>
- Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.
- Biermann, M., Bohus, M., Gilbert, P., Vonderlin, R., Cornelisse, S., Osen, B., Graser, J., Brüne, M., Ebert, A., Kleindienst, N., & Lyssenko, L. (2020). Psychometric properties of the German version of the Forms of Self-Criticizing/Attacking and Self-Reassuring Scale (FSCRS). *Psychological Assessment, 33*(1), 97–110. <https://doi.org/10.1037/pas0000956>
- Blatt, S. J., Quinlan, D. M., Chevron, E. S., McDonald, C., & Zuroff, D. (1982). Dependency and self-criticism: Psychological dimensions of depression. *Journal of Consulting and Clinical Psychology, 50*, 113–124. <https://doi.org/10.1037/0022-006X.50.1.113>
- Blatt, S. J., Quinlan, D. M., Pilkonis, P. A., & Shea, M. T. (1995). Impact of perfectionism and need for approval on the brief treatment of depression: The National Institute of Mental Health Treatment of Depression Collaborative Research Program revisited. *Journal of Consulting and Clinical Psychology, 63*(1), 125–132. <https://doi.org/10.1037/0022-006X.63.1.125>
- Blatt, S. J., & Zuroff, D. C. (1992). Interpersonal relatedness and self definition: Two prototypes for depression. *Clinical Psychology Review, 12*, 527–562. [https://doi.org/10.1016/0272-7358\(92\)90070-O](https://doi.org/10.1016/0272-7358(92)90070-O)
- Bluth, K., Campo, R. A., Futch, W. S., & Gaylord, S. A. (2017). Age and gender differences in the associations of self-compassion and emotional well-being in a large adolescent sample. *Journal of Youth and Adolescence, 46*(4), 840–853. <https://doi.org/10.1007/s10964-016-0567-2>
- Boellinghaus, I., Jones, F. W., & Hutton, J. (2014). The role of mindfulness and loving-kindness meditation in cultivating self-compassion and other-focused concern in health care professionals. *Mindfulness, 5*(2), 129–138.
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. Chichester: John Wiley and Sons Ltd.
- Bowlby, J. (1969). *Attachment and loss: Vol. 1. Attachment* (2nd ed.). New York: Basic Books.
- Bowlby, J. (1973). *Attachment and loss: Separation, anxiety and anger*. New York: Basic Books.
- Braehler, C., Gumley, A., Harper, J., Wallace, S., Norrie, J., & Gilbert, P. (2013). Exploring change processes in compassion focused therapy in psychosis: Results of a feasibility randomized controlled trial. *British Journal of Clinical Psychology, 52*, 199–214. <https://doi.org/10.1111/bjc.12009>
- Cantazaro, A., & Wei, M. (2010). Adult attachment, dependence, self-criticism, and depressive symptoms: A test of a mediational model. *Journal of Personality, 78*(4), 1135–1162. <https://doi.org/10.1111/j.1467-6494.2010.00645>
- Carr, E. R., Szymanski, D. M., Taha, F., West, L. M., & Kaslow, N. J. (2014). Understanding the link between multiple oppressions and depression among African American women: The role of internalization. *Psychology of Women Quarterly, 38*(2), 233–245. <https://doi.org/10.1177/0361684313499900>
- Castilho, P., Pinto-Gouveia, J., & Duarte, J. (2015). Exploring self-criticism: Confirmatory factor analysis of the FSCRS in clinical and nonclinical samples. *Clinical Psychology & Psychotherapy, 22*(2), 153–164. <https://doi.org/10.1002/cpp.1881>
- Castilho, P., Pinto-Gouveia, J., & Duarte, J. (2017). Two forms of self-criticism mediate differently the shame–psychopathological symptoms link. *Psychology and Psychotherapy: Theory, Research and Practice, 90*(1), 44–54. <https://doi.org/10.1111/papt.12094>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, New Jersey, USA: Laurence Erlbaum Associates.
- Connor, K. M., Davidson, J. R., Churchill, L. E., Sherwood, A., Weisler, R. H., & Foa, E. (2000). Psychometric properties of the Social Phobia Inventory (SPIN): New self-rating scale. *The British Journal of Psychiatry, 176*(4), 379–386. <https://doi.org/10.1192/bjp.176.4.379>
- Cornish, M. A., & Wade, N. G. (2015). Working through past wrongdoing: Examination of a self-forgiveness counselling intervention. *Journal of Counselling Psychology, 62*(3), 521–528. <https://doi.org/10.1037/cou0000080>
- Cox, B. J., Fleet, C., & Stein, M. B. (2004). Self-criticism and social phobia in the US national comorbidity survey. *Journal of Affective Disorders, 82*, 227–234. <https://doi.org/10.1016/j.jad.2003.12.012>
- Cox, B. J., MacPherson, P. S. R., Enns, M. W., & McWilliams, L. A. (2004). Neuroticism and self-criticism associated with posttraumatic stress disorder in a nationally representative sample. *Behaviour Research and Therapy, 42*, 105–114. [https://doi.org/10.1016/S0005-7967\(03\)00105-0](https://doi.org/10.1016/S0005-7967(03)00105-0)
- Craig, C., Hiskey, S., & Spector, A. (2020). Compassion focused therapy: A systematic review of its effectiveness and acceptability in clinical populations. *Expert Review of Neurotherapeutics, 20*(4), 385–400. <https://doi.org/10.1080/14737175.2020.1746184>
- de Graaf, L. E., Roelofs, J., & Huibers, M. J. H. (2009). Measuring dysfunctional attitudes in the general population: The Dysfunctional Attitude Scale (form A) revised. *Cognitive Therapy and Research, 33*(4), 345–355. <https://doi.org/10.1007/s10608-009-9229-y>

- Duarte, C., Pinto-Gouveia, J., & Stubbs, R. J. (2017). Compassionate attention and regulation of eating behaviour: A pilot study of a brief low-intensity intervention for binge eating. *Clinical Psychology & Psychotherapy*, 24(6), O1437–O1447. <https://doi.org/10.1002/cpp.2094>
- Dundas, I., Binder, P. E., Hansen, T. G., & Stige, S. H. (2017). Does a short self-compassion intervention for students increase healthy self-regulation? A randomized control trial. *Scandinavian Journal of Psychology*, 58(5), 443–450. <https://doi.org/10.1111/sjop.12385>
- Dunkley, D. M., & Grilo, C. M. (2007). Self-criticism, low self-esteem, depressive symptoms, and over-evaluation of shape and weight in binge eating disorder patients. *Behaviour Research and Therapy*, 45, 39–149. <https://doi.org/10.1016/j.brat.2006.01.017>
- Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455–463. <https://doi.org/10.1111/j.0006-341X.2000.00455.x>
- Enns, M. W., Cox, B. J., & Pidlubny, S. R. (2002). Group cognitive behaviour therapy for residual depression: Effectiveness and predictors of response. *Cognitive Behaviour Therapy*, 31(1), 31–40. <https://doi.org/10.1080/16506070252823643>
- Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, 16, 363–370. [https://doi.org/10.1002/1098-108X\(199412\)16:4<363::AID-EAT2260160405>3.0.CO;2-%23](https://doi.org/10.1002/1098-108X(199412)16:4<363::AID-EAT2260160405>3.0.CO;2-%23)
- Fazaa, N., & Page, S. (2009). Personality style and impulsivity as determinants of suicidal subgroups. *Archives of Suicide Research*, 13(1), 31–45. <https://doi.org/10.1080/13811110802572122>
- Feliu-Soler, A., Pascual, J. C., Elices, M., Martín-Blanco, A., Carmona, C., Cebolla, A., Simón, V., & Soler, J. (2017). Fostering self-compassion and loving-kindness in patients with borderline personality disorder: A randomized pilot study. *Clinical Psychology & Psychotherapy*, 24(1), 278–286. <https://doi.org/10.1002/cpp.2000>
- Ferrari, M., Hunt, C., Harrysunker, A., Abbott, M. J., Beath, A. P., & Einstein, D. A. (2019). Self-compassion interventions and psychosocial outcomes: A meta-analysis of RCTs. *Mindfulness*, 10(8), 1455–1473. <https://doi.org/10.1007/s12671-019-01134-6>
- Ferrari M., Yap K., Scott N., Einstein D. A., Ciarrochi J. (2018). Self-compassion moderates the perfectionism and depression link in both adolescence and adulthood. *PLOS ONE*, 13, (2), e0192022. <https://doi.org/10.1371/journal.pone.0192022>
- Finch, B. K., Kolody, B., & Vega, W. A. (2000). Perceived discrimination and depression among Mexican-origin adults in California. *Journal of Health and Social Behavior*, 41, 295–313. <https://doi.org/10.2307/2676322>
- First, M., Spitzer, R., Gibbon, M., Williams, J., & Benjamin, L. (1994). *Structured clinical interview for DSM-IV axis II personality disorders (SCID II)*. New York: Biometric Research Department.
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1997). *Structured clinical interview for DSM-IV axis I disorders (SCID I)*. New York: Biometric Research Department.
- Fisher, M. L., & Exline, J. J. (2006). Self-forgiveness versus excusing: The roles of remorse, effort, and acceptance of responsibility. *Self and Identity*, 5, 127–146. <https://doi.org/10.1080/15298860600586123>
- Gale, C., Gilbert, P., Read, N., & Goss, K. (2014). An evaluation of the impact of introducing compassion focused therapy to a standard treatment programme for people with eating disorders. *Clinical Psychology & Psychotherapy*, 21(1), 1–12. <https://doi.org/10.1002/cpp.1806>
- Germer, C. K. (2009). *The mindful path to self-compassion: Freeing yourself from destructive thoughts and emotions*. New York, NY: Guilford Press.
- Gilbert, P. (2009). Introducing compassion-focused therapy. *Advances in Psychiatric Treatment*, 15(3), 199–208. <https://doi.org/10.1192/apt.bp.107.005264>
- Gilbert, P. (2010). *Compassion focused therapy: Distinctive features*. London: Routledge.
- Gilbert, P. (2014). The origins and nature of compassion focused therapy. *British Journal of Clinical Psychology*, 53, 6–41. <https://doi.org/10.1111/bjc.12043>
- Gilbert, P., Basran, J., MacArthur, M., & Kirby, J. N. (2019). Differences in the semantics of prosocial words: An exploration of compassion and kindness. *Mindfulness*, 10(11), 2259–2271. <https://doi.org/10.1007/s12671-019-01191-x>
- Gilbert, P., Clarke, M., Hempel, S., Miles, J. N., & Irons, C. (2004). Criticizing and reassuring oneself: An exploration of forms, styles and reasons in female students. *British Journal of Clinical Psychology*, 43(1), 31–50. <https://doi.org/10.1348/014466504772812959>
- Gilbert, P., McEwan, K., Irons, C., Bhundia, R., Christie, R., Broomhead, C., & Rockliff, H. (2010). Self-harm in a mixed clinical population: The roles of self-criticism, shame, and social rank. *British Journal of Clinical Psychology*, 49, 563–576. <https://doi.org/10.1348/014466509X479771>
- Gilbert, P., & Miles, J. N. (2000). Sensitivity to social put-down: It's relationship to perceptions of social rank, shame, social anxiety, depression, anger and self-other blame. *Personality and Individual Differences*, 29(4), 757–774. [https://doi.org/10.1016/S0191-8869\(99\)00230-5](https://doi.org/10.1016/S0191-8869(99)00230-5)
- Gilbert, P., & Procter, S. (2006). Compassionate mind training for people with high shame and self-criticism: Overview and pilot study of a group therapy approach. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice*, 13(6), 353–379. <https://doi.org/10.1002/cpp.507>
- Glassman, L. H., Weierich, M. R., Hooley, J. M., Deliberto, T. L., & Nock, M. K. (2007). Child maltreatment, non-suicidal self-injury, and the mediating role of self-criticism. *Behaviour Research and Therapy*, 45(10), 2483–2490. <https://doi.org/10.1016/j.brat.2007.04.002>
- Goss, K. (2011). *The compassionate-mind guide to ending overeating: Using compassion-focused therapy to overcome bingeing and disordered eating*. Oakland, California: New Harbinger.
- Greenberg, L. (2002). *Emotion-focused therapy: Coaching clients to work through feelings*. Washington, DC: American Psychological Association.
- Halamová, J., Kanovský, M., Gilbert, P., Troop, N. A., Zuroff, D. C., Hermanto, N., Petrocchi, N., Sommers-Spijkerman, M., Kirby, J. N., Shahar, B., & Krieger, T. (2018). The factor structure of the forms of self-criticising/attacking & self-reassuring scale in thirteen distinct populations. *Journal of Psychopathology and Behavioral Assessment*, 40(4), 736–751. <https://doi.org/10.1007/s10862-018-9686-2>
- Halamová, J., Kanovský, M., & Pacúchová, M. (2017). Robust psychometric analysis and factor structure of the forms of self-criticizing/attacking and self-reassuring scale. *Československá Psychologie*, 61(4), 331–349.
- Halamová, J., Kanovský, M., Varšová, K., & Kupelí, N. (2018). Randomised controlled trial of the new short-term online emotion focused training for self-compassion and self-protection in a nonclinical sample. *Current Psychology*, 40, 1–11. <https://doi.org/10.1007/s12144-018-9933-4>
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando: Academic Press.
- Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods*, 9(4), 426–445. <https://doi.org/10.1037/1082-989X.9.4.426>
- Hempel, S., Miles, J. N., Booth, M. J., Wang, Z., Morton, S. C., & Shekelle, P. G. (2013). Risk of bias: A simulation study of power to detect study-level moderator effects in meta-analysis. *Systematic Reviews*, 2(1), 107. <https://doi.org/10.1186/2046-4053-2-107>
- Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savović, J., Schulz, K. F., Weeks, L., & Sterne, J. A. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *British Medical Journal*, 343, d5928. <https://doi.org/10.1136/bmj.d5928>
- Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.) (2020). *Cochrane handbook for systematic reviews of interventions, Version 6.1*. New Jersey USA: John Wiley & Sons.

- Hofmann, S. G., Grossman, P., & Hinton, D. E. (2011). Lovingkindness and compassion meditation: Potential for psychological interventions. *Clinical Psychology Review*, 31, 1126–1132. <https://doi.org/10.1016/j.cpr.2011.07.003>
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70.
- Homan, K. J. (2016). Self-compassion and psychological well-being in older adults. *Journal of Adult Development*, 23(2), 111–119. <https://doi.org/10.1007/s10804-016-9227-8>
- Hutton, P., Kelly, J., Lowens, I., Taylor, P. J., & Tai, S. (2013). Self-attacking and self-reassurance in persecutory delusions: A comparison of healthy, depressed and paranoid individuals. *Psychiatry Research*, 205(1–2), 127–136. <https://doi.org/10.1016/j.psychres.2012.08.010>
- Johnson, S. B., Goodnight, B. L., Zhang, H., Daboin, I., Patterson, B., & Kaslow, N. J. (2018). Compassion-based meditation in African Americans: Self-criticism mediates changes in depression. *Suicide and Life-Threatening Behavior*, 48(2), 160–168. <https://doi.org/10.1111/sltb.12347>
- Kabat-Zinn, J. (1990). *Full catastrophe living: Using the wisdom of your body and mind to face stress, pain, and illness*. New York: Dell Publishing.
- Kannan, D., & Levitt, H. M. (2013). A review of client self-criticism in psychotherapy. *Journal of Psychotherapy Integration*, 23, 166–178. <https://doi.org/10.1037/a0032355>
- Kelly, A. C., & Carter, J. C. (2015). Self-compassion training for binge eating disorder: A pilot randomized controlled trial. *Psychology and Psychotherapy: Theory, Research and Practice*, 88(3), 285–303. <https://doi.org/10.1111/papt.12044>
- Kelly, A. C., & Leybman, M. J. (2012). Calming your eating disorder voice with compassion-focused therapy. Unpublished manual.
- Kelly, A. C., Wisniewski, L., Martin-Wagar, C., & Hoffman, E. (2017). Group-based compassion-focused therapy as an adjunct to outpatient treatment for eating disorders: A pilot randomized controlled trial. *Clinical Psychology & Psychotherapy*, 24(2), 475–487. <https://doi.org/10.1002/cpp.2018>
- Kelman, A. R., Evare, B. S., Barrera, A. Z., Muñoz, R. F., & Gilbert, P. (2018). A proof-of-concept pilot randomized comparative trial of brief internet-based compassionate mind training and cognitive-behavioral therapy for perinatal and intending to become pregnant women. *Clinical Psychology & Psychotherapy*, 25(4), 608–619. <https://doi.org/10.1002/cpp.2185>
- Kemeny, M. E., Foltz, C., Cavanagh, J. F., Cullen, M., Giese-Davis, J., Jennings, P., Rosenberg, E. L., Gillath, O., Shaver, P. R., Wallace, B. A., & Ekman, P. (2012). Contemplative/emotion training reduces negative emotional behavior and promotes prosocial responses. *Emotion*, 12(2), 338–350. <https://doi.org/10.1037/a0026118>
- Kim, J. J., Henderson, T., Best, T., Cunnington, R., & Kirby, J. N. (2020). Neural and self-report markers of reassurance: A generalized additive modelling approach. *Frontiers in Psychiatry*, 11, 566141. <https://doi.org/10.3389/fpsy.2020.566141>
- Kim, J. J., Kent, K. M., Cunnington, R., Gilbert, P., & Kirby, J. N. (2020). Attachment styles modulate neural markers of threat and imagery when engaging in self-criticism. *Scientific Reports*, 10(1), 1–10. <https://doi.org/10.1038/s41598-020-70772->
- Kirby, J. N., & Baldwin, S. (2018). A randomized micro-trial of a loving-kindness meditation to help parents respond to difficult child behaviour vignettes. *Journal of Child and Family Studies*, 27(5), 1614–1628. <https://doi.org/10.1007/s10826-017-0989-9>
- Kirby, J. N., Day, J., & Sagar, V. (2019). The ‘flow’ of compassion: A meta-analysis of the fears of compassion scales and psychological functioning. *Clinical Psychology Review*, 70, 26–39. <https://doi.org/10.1016/j.cpr.2019.03.001>
- Kirby, J. N., & Laczko, D. (2017). A randomized micro-trial of a loving-kindness meditation for young adults living at home with their parents. *Journal of Child and Family Studies*, 26(7), 1888–1899. <https://doi.org/10.1007/s10826-017-0692-x>
- Kirby, J. N., Tellegen, C. L., & Steindl, S. R. (2017). A meta-analysis of compassion-based interventions: Current state of knowledge and future directions. *Behavior Therapy*, 48(6), 778–792. <https://doi.org/10.1016/j.beth.2017.06.003>
- Kitayama, S., Markus, H. R., Matsumoto, H., & Norasakkunkit, V. (1997). Individual and collective processes in the construction of the self: Self-enhancement in the United States and self-criticism in Japan. *Journal of Personality and Social Psychology*, 72(6), 1245–1267. <https://doi.org/10.1037/0022-3514.72.6.1245>
- Klomek, A. B., Orbach, I., Sher, L., Sommerfeld, E., Diller, R., Apter, A., Shahar, G., & Zalsman, G. (2008). Quality of depression among suicidal inpatient youth. *Archives of Suicide Research*, 12(2), 133–140. <https://doi.org/10.1080/13811110701857160>
- Kopala-Sibley, D. C., Zuroff, D. C., Russell, J. J., Moskowitz, D. S., & Paris, J. (2012). Understanding heterogeneity in borderline personality disorder: Differences in affective reactivity explained by the traits of dependency and self-criticism. *Journal of Abnormal Psychology*, 121, 680–691. <https://doi.org/10.1037/a0028513>
- Krieger, T., Reber, F., von Glutz, B., Urech, A., Moser, C. T., Schulz, A., & Berger, T. (2019). An internet-based compassion-focused intervention for increased self-criticism: A randomized controlled trial. *Behavior Therapy*, 50(2), 430–445. <https://doi.org/10.1016/j.beth.2018.08.003>
- Kupeli, N., Chilcot, J., Schmidt, U. H., Campbell, I. C., & Troop, N. A. (2013). A confirmatory factor analysis and validation of the forms of self-criticism/reassurance scale. *British Journal of Clinical Psychology*, 52, 12–25. <https://doi.org/10.1111/j.2044-8260.2012.02042.x>
- Lagakos, S. W. (2006). The challenge of subgroup analyses-reporting without distorting. *New England Journal of Medicine*, 354(16), 1667–1669.
- Lamers, S. M. A., Westerhof, G. J., Bohlmeijer, E. T., ten Klooster, P. M., & Keyes, C. L. (2011). Evaluating the psychometric properties of the mental health continuum-short form (MHC-SF). *Journal of Clinical Psychology*, 67(1), 99–110. <https://doi.org/10.1002/jclp.20741>
- Leaviss, J., & Uttley, L. (2015). Psychotherapeutic benefits of compassion-focused therapy: An early systematic review. *Psychological Medicine*, 45(5), 927–945. <https://doi.org/10.1017/S0033291714002141>
- Longe, O., Maratos, F. A., Gilbert, P., Evans, G., Volker, F., Rockliff, H., & Rippon, G. (2010). Having a word with yourself: Neural correlates of self-criticism and self-reassurance. *NeuroImage*, 49(2), 1849–1856. <https://doi.org/10.1016/j.neuroimage.2009.09.019>
- Luberto, C. M., Shinday, N., Song, R., Philpotts, L. L., Park, E. R., Fricchione, G. L., & Yeh, G. Y. (2018). A systematic review and meta-analysis of the effects of meditation on empathy, compassion, and prosocial behaviors. *Mindfulness*, 9(3), 708–724. <https://doi.org/10.1007/s12671-017-0841-8>
- Lucre, K. M., & Corten, N. (2013). An exploration of group compassion-focused therapy for personality disorder. *Psychology and Psychotherapy: Theory, Research and Practice*, 86(4), 387–400. <https://doi.org/10.1111/j.2044-8341.2012.02068.x>
- Lutz, J., Berry, M. P., Napadow, V., Germer, C., Pollak, S., Gardiner, P., Edwards, R. R., Desbordes, G., & Schuman-Olivier, Z. (2020). Neural activations during self-related processing in patients with chronic pain and effects of a brief self-compassion training—A pilot study. *Psychiatry Research: Neuroimaging*, 304, 111155. <https://doi.org/10.1016/j.psychres.2020.111155>
- Luyten, P., Sabbe, B., Blatt, S. J., Meganck, S., Jansen, B., de Grave, C., Maes, F., & Corveleyn, J. (2007). Dependency and self-criticism: Relationship with major depressive disorder, severity of depression, and clinical presentation. *Depression and Anxiety*, 24, 586–596. <https://doi.org/10.1002/da.20272>
- Matos, M., Duarte, C., Duarte, J., Pinto-Gouveia, J., Petrocchi, N., Basran, J., & Gilbert, P. (2017). Psychological and physiological effects of compassionate mind training: A pilot randomised controlled study. *Mindfulness*, 8(6), 1699–1712. <https://doi.org/10.1007/s12671-017-0745-7>

- Matthews, D. D., Hammond, W. P., Nuru-Jeter, A., Cole-Lewis, Y., & Melvin, T. (2013). Racial discrimination and depressive symptoms among African-American men: The mediating and moderating roles of masculine self-reliance and John Henryism. *Psychology of Men & Masculinity*, 14(1), 35–46. <https://doi.org/10.1037/a0028436>
- Mayhew, S. L., & Gilbert, P. (2008). Compassionate mind training with people who hear malevolent voices: A case series report. *Clinical Psychology & Psychotherapy*, 15, 113–138. <https://doi.org/10.1002/cpp.566>
- Miranda, J., Nakamura, R., & Bernal, G. (2003). Including ethnic minorities in mental health intervention research: A practical approach to a long-standing problem. *Culture, Medicine and Psychiatry*, 27(4), 467–486. <https://doi.org/10.1023/B:MEDI.0000005484.26741.79>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Reprint—Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Physical Therapy*, 89(9), 873–880. <https://doi.org/10.1093/ptj/89.9.873>
- Mosewich, A. D., Crocker, P. R., Kowalski, K. C., & DeLongis, A. (2013). Applying self-compassion in sport: An intervention with women athletes. *Journal of Sport and Exercise Psychology*, 35(5), 514–524. <https://doi.org/10.1123/jsep.35.5.514>
- Naismith, I., Zarate Guerrero, S., & Feigenbaum, J. (2019). Abuse, invalidation, and lack of early warmth show distinct relationships with self-criticism, self-compassion, and fear of self-compassion in personality disorder. *Clinical Psychology & Psychotherapy*, 26(3), 350–361. <https://doi.org/10.1002/cpp.2357>
- National Institute for Health and Care Excellence. (2012). Methods for the development of NICE public health guidance (3<sup>rd</sup> ed). Retrieved April, 2020, from <https://www.nice.org.uk/process/pmg4/chapter/appendix-f-quality-appraisal-checklist-quantitative-intervention-studies>
- Neff, K. (2003b). Self-compassion: An alternative conceptualization of a healthy attitude toward oneself. *Self and Identity*, 2, 85–101. <https://doi.org/10.1080/15298860309032>
- Neff, K. D. (2003a). The development and validation of a scale to measure self-compassion. *Self and Identity*, 2(3), 223–250. <https://doi.org/10.1080/15298860309027>
- Neff, K. D., & Germer, C. K. (2013). A pilot study and randomized controlled trial of the mindful self-compassion program. *Journal of Clinical Psychology*, 69(1), 28–44. <https://doi.org/10.1002/jclp.21923>
- Noh, S., Beiser, M., Kaspar, V., Hou, F., & Rummens, J. (1999). Perceived racial discrimination, depression, and coping: A study of Southeast Asian refugees in Canada. *Journal of Health and Social Behavior*, 40, 193–207. <https://doi.org/10.2307/2676348>
- Ozawa-de Silva, B., & Negi, L. T. (2013). Cognitively-based compassion training (CBCT)—Protocol and key concepts. In T. Singer & M. Bolz (Eds.), *Compassion: Bridging theory and practice: A multimedia book* (pp. 417–438). Leipzig, Germany: Max-Planck Institute.
- Palmeira, L., Pinto-Gouveia, J., & Cunha, M. (2017). Exploring the efficacy of an acceptance, mindfulness & compassionate-based group intervention for women struggling with their weight (Kg-Free): A randomized controlled trial. *Appetite*, 112, 107–116. <https://doi.org/10.1016/j.appet.2017.01.027>
- Perry, B. L., Harp, K. L., & Oser, C. B. (2013). Racial and gender discrimination in the stress process: Implications for African American women's health and well-being. *Sociological Perspectives*, 56(1), 25–48. <https://doi.org/10.1525/sop.2012.56.1.25>
- 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. <https://doi.org/10.1002/cpp.702>
- Rector, N. A., Bagby, R. M., Segal, Z. V., Joffe, R. T., & Levitt, A. (2000). Self-criticism and dependency in depressed patients treated with cognitive therapy or pharmacotherapy. *Cognitive Therapy and Research*, 24, 571–584. <https://doi.org/10.1023/A:1005566112869>
- Rosenthal, R. (1991). *Meta-analytic procedures for social research*. California: Sage Publications.
- Salzberg, S. (1995). *Loving kindness: The revolutionary art of happiness*. Boston, MA: Shambala.
- Santorelli, S., & Kabat-Zinn, J. (2009). *Mindfulness-based stress reduction (MBSR) professional education and training: MBSR curriculum and supporting materials*. Worcester, MA: University of Massachusetts Medical School, Dept. of Medicine.
- Shahar, B., Szepsenwol, O., Zilcha-Mano, S., Haim, N., Zamir, O., Levi-Yeshuvi, S., & Levit-Binnun, N. (2015). A wait-list randomized controlled trial of loving-kindness meditation programme for self-criticism. *Clinical Psychology & Psychotherapy*, 22(4), 346–356. <https://doi.org/10.1002/cpp.1893>
- Shahar, G. (2015). *Erosion: The psychopathology of self-criticism*. New York: Oxford University Press.
- Shahar, G., Blatt, S. J., Zuroff, D. C., & Pilkonis, P. A. (2003). Role of perfectionism and personality disorder features in response to brief treatment for depression. *Journal of Consulting and Clinical Psychology*, 71(3), 629–633. <https://doi.org/10.1037/0022-006X.71.3.629>
- Sommers-Spijkerman, M., Trompetter, H., ten Klooster, P., Schreurs, K., Gilbert, P., & Bohlmeijer, E. (2018). Development and validation of the forms of Self-Criticizing/Attacking and Self-Reassuring Scale—Short form. *Psychological Assessment*, 30(6), 729–743. <https://doi.org/10.1037/pas0000514>
- Sommers-Spijkerman, M. P. J., Trompetter, H. R., Schreurs, K. M. G., & Bohlmeijer, E. T. (2018). Compassion-focused therapy as guided self-help for enhancing public mental health: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 86(2), 101–115. <https://doi.org/10.1037/ccp0000268>
- Spitzer, R. L., Kroenke, K., & Williams, J. B. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. *Journal of the American Medical Association*, 282, 1737–1744. <https://doi.org/10.1001/jama.282.18.1737>
- Stevenson, J., Mattiske, J. K., & Nixon, R. D. (2019). The effect of a brief online self-compassion versus cognitive restructuring intervention on trait social anxiety. *Behaviour Research and Therapy*, 123, 103492. <https://doi.org/10.1016/j.brat.2019.103492>
- Thompson, R. S., & Zuroff, D. C. (2004). The Levels of Self-Criticism Scale: Comparative self-criticism and internalized self-criticism. *Personality and Individual Differences*, 36, 419–430. [https://doi.org/10.1016/S0191-8869\(03\)00106-5](https://doi.org/10.1016/S0191-8869(03)00106-5)
- Van de Brink, E., & Koster, F. (2015). *Mindfulness-based compassionate living: A new training programme to deepen mindfulness with heartfulness*. London: Routledge.
- Verplanken, B., Friborg, O., Wang, C. E., Traminow, D., & Woolf, K. (2007). Mental habits: Metacognitive reflection on negative self thinking. *Journal of Personality and Social Psychology*, 92, 526–554. <https://doi.org/10.1037/0022-3514.92.3.526>
- Wallace, B. A. (1999). *Boundless heart, the cultivation of the four immeasurables*. Boston: Snow Lion Publications.
- Wallmark, E., Safarzadeh, K., Daukantaitė, D., & Maddux, R. E. (2013). Promoting altruism through meditation: An 8-week randomized controlled pilot study. *Mindfulness*, 4, 223–234. <https://doi.org/10.1007/s12671-012-0115-4>
- Wang, R., & Ware, J. H. (2013). Detecting moderator effects using subgroup analyses. *Prevention Science*, 14(2), 111–120. <https://doi.org/10.1007/s11121-011-0221-x>
- Weissman, A. N., & Beck, A. T. (1978). Development and validation of the Dysfunctional Attitude Scale: A preliminary investigation. Paper presented at the 86<sup>th</sup> Annual Convention of the American Psychological Association, Toronto, Canada.
- Werner, A. M., Tibubos, A. N., Rohrmann, S., & Reiss, N. (2019). The clinical trait self-criticism and its relation to psychopathology: A systematic review—Update. *Journal of Affective Disorders*, 246, 530–547. <https://doi.org/10.1016/j.jad.2018.12.069>



- Whelton, W. J., & Greenberg, L. S. (2005). Emotion in self-criticism. *Personality and Individual Differences*, 38, 1583–1595. <https://doi.org/10.1016/j.paid.2004.09.024>
- Wilson, A. C., Mackintosh, K., Power, K., & Chan, S. W. (2019). Effectiveness of self-compassion related therapies: A systematic review and meta-analysis. *Mindfulness*, 10(6), 979–995. <https://doi.org/10.1007/s12671-018-1037-6>
- World Health Organization. (2004). *ICD-10: International statistical classification of diseases and related health problems: Tenth revision* (2nd ed.). Geneva: World Health Organization.
- Xavier, A., Cunha, M., & Pinto Gouveia, J. (2015). Deliberate self-harm in adolescence: The impact of childhood experiences, negative affect and fears of compassion. *Revista de Psicopatologia y Psicología Clínica*, 20(1), 41–49. <https://doi.org/10.5944/rppc.vol.1.num.1.2015.14407>
- 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(6), 1136–1152. <https://doi.org/10.1007/s12671-018-1066-1>
- 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(5), 499–520. <https://doi.org/10.1080/15298868.2015.1029966>
- Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67(6), 361–370. <https://doi.org/10.1111/j.1600-0447.1983.tb09716.x>
- Zuroff, D. C., Blatt, S. J., Sotsky, S. M., Krupnick, J. L., Martin, D. J., Sanislow, C. A. III, & Simmens, S. (2000). Relation of therapeutic alliance and perfectionism to outcome in brief outpatient treatment of depression. *Journal of Consulting and Clinical Psychology*, 68(1), 114–124. <https://doi.org/10.1037/0022-006X.68.1.114>
- Zuroff, D. C., Moskowitz, D. S., & Côté, S. (1999). Dependency, self-criticism, interpersonal behaviour and affect: Evolutionary perspectives. *British Journal of Clinical Psychology*, 38(3), 231–250. <https://doi.org/10.1348/014466599162827>

**How to cite this article:** Wakelin KE, Perman G, Simonds LM. Effectiveness of self-compassion-related interventions for reducing self-criticism: A systematic review and meta-analysis. *Clin Psychol Psychother*. 2021;1–25. <https://doi.org/10.1002/cpp.2586>