



Systematic Review

The Effectiveness of Compassion Focused Therapy for the Three Flows of Compassion, Self-Criticism, and Shame in Clinical Populations: A Systematic Review

Naomi Brown ^{1,*}  and Katie Ashcroft ² ¹ Berkshire Healthcare Foundation Trust, Bracknell RG12 7FR, UK² Department of Psychology, Royal Holloway University of London, Egham TW20 0EX, UK; katie.ashcroft@rhul.ac.uk

* Correspondence: naomic.brown@berkshire.nhs.uk

Abstract

Compassion Focused therapy (CFT) is designed to reduce shame (internal and external) and self-criticism while enhancing the three flows of compassion (compassion to others, from others, and for the self). This systematic review evaluated the effectiveness of CFT on these core theoretical constructs in adult clinical populations. A systematic search of three databases (2000–2024) identified 21 studies ($N = 450$) meeting the inclusion criteria. The studies were narratively synthesised, and quality was assessed using the EPHPP tool. Consistent improvements in self-compassion ($g = 0.23$ – 4.14) and reductions in self-criticism ($g = 0.29$ – 1.56) were reported. Reductions in external shame were also observed ($g = 0.54$ – 1.22), though this outcome was examined in fewer studies. Limited and inconsistent evidence was found for internal shame and interpersonal compassion flows (compassion to and from others), with only a small number of low- to moderate-quality studies addressing these outcomes. Follow-up effects were rarely assessed, and comparator groups were limited. Most interventions were group-based and of variable methodological quality, with frequent selection bias, small sample sizes, and limited demographic diversity. Overall, CFT shows promise for targeting self-directed processes in clinical populations, though stronger evidence is needed to understand its effects on relational components of compassion. Future research should adopt standardised measures, improve methodological rigour, and recruit more diverse samples.

Keywords: clinical populations; compassion-focused therapy; compassion flows; self-compassion; self-criticism; shame; systematic review



Academic Editor: Andrew Soundy

Received: 30 May 2025

Revised: 24 July 2025

Accepted: 27 July 2025

Published: 29 July 2025

Citation: Brown, N., & Ashcroft, K. (2025). The Effectiveness of Compassion Focused Therapy for the Three Flows of Compassion, Self-Criticism, and Shame in Clinical Populations: A Systematic Review. *Behavioral Sciences*, 15(8), 1031. <https://doi.org/10.3390/bs15081031>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Compassion Focused therapy (CFT), developed by Gilbert (2000), is a psychological intervention that addresses mental health difficulties through compassion cultivation. Drawing from evolutionary psychology and neuroscience, CFT targets common factors in mental health conditions, particularly self-criticism and shame, through fostering compassionate behaviour (Gilbert & Choden, 2013; Gilbert, 2014, 2020).

Within CFT, self-criticism and shame are understood as barriers to wellbeing (Gilbert & Procter, 2006). Self-criticism involves harsh self-judgment, while shame encompasses feelings of unworthiness. Shame can be directed internally (shame directed towards oneself) or externally (perceiving others as shaming; Callow et al., 2021; Gilbert, 2014; Lee et al., 2001). CFT theory suggests that both internal and external shame activate the

threat system, creating self-protective responses that paradoxically maintain psychological distress (Gilbert & Simos, 2022). CFT interventions theoretically engage with internal shame through developing self-compassion and a kinder internal relationship, while addressing external shame by fostering secure connections and receptivity to compassion from others (Gilbert & Simos, 2022). Persistent self-judgment is viewed as a barrier to self-compassion and emotional regulation, with self-criticism and shame directly contributing to and maintaining psychological distress (Gilbert & Procter, 2006; Gilbert & Simos, 2022).

A CFT approach defines compassion as sensitivity and commitment to alleviating suffering (Gilbert, 2014). CFT theory proposes three emotion regulation systems: threat (safety-seeking), drive (resource-seeking), and soothing (affiliation-seeking). The model suggests that cultivating a compassionate motivation, balances these systems by activating the soothing system (Gilbert & Simos, 2022). The soothing system is then used to counteract excessive threat activation and behaviour motivated by threat (Gilbert, 2014, 2020).

Compassion in CFT operates through three distinct ‘flows’, each serving vital yet complementary functions (Gilbert, 2014). These include (a) Compassion for Others: nurturing social connections through caring behaviour directed toward others; (b) Compassion from Others: developing receptivity to care and support from others, which affirms one’s worthiness of compassion; and (c) Self-Compassion: cultivating emotional resilience through treating oneself with kindness and understanding. These flows work together to promote psychological wellbeing by addressing both interpersonal and intrapersonal dimensions of compassion (Gilbert, 2020).

1.1. Previous Reviews and Current Gaps

Early systematic review evidence (Leaviss & Uttley, 2015) demonstrated CFT’s efficacy for reducing self-criticism and improving self-compassion in clinical populations. However, this review excluded key theoretical mechanisms like shame and interpersonal compassion flows. Later reviews (Kirby et al., 2017) examined multiple ‘compassion-based’ interventions together, limiting conclusions about CFT’s unique therapeutic mechanisms. This trend of grouping compassion interventions persists in recent reviews (Wakelin et al., 2021; Wilson et al., 2018), despite concerns that this approach obscures the unique mechanisms of each intervention (Kirby & Gilbert, 2019). Recent CFT-specific reviews in clinical populations (Craig et al., 2020; Millard et al., 2023) have focused on general symptoms rather than CFT’s theoretical mechanisms of change, such as compassion flows, self-criticism, and shame (Gilbert & Simos, 2022). This gap limits our understanding of how CFT achieves its therapeutic effects. The current review addresses these gaps by synthesising evidence for CFT’s impact on its core theoretical constructs in clinical populations. Given the limited RCT evidence for shame and compassion flows, the review included both RCTs and other study designs to provide a comprehensive evaluation.

1.2. Aims

This systematic review evaluates the effectiveness of CFT for clinical populations, focusing on three key domains: the three flows of compassion, self-criticism, and shame.

Primary research questions examine CFT’s effectiveness in changing

1. The three flows of compassion (self-compassion and interpersonal compassion: compassion from others, and compassion to others);
2. Self-criticism;
3. Shame (internal and external).

The review examines changes in these outcomes from baseline to post-intervention and follow-up, analysing published and unpublished CFT studies from 2000 to 2024 through narrative synthesis.

2. Materials and Methods

This systematic review followed the PRISMA guidelines (Page et al., 2021) and was pre-registered with PROSPERO (CRD42023441154).

2.1. Eligibility Criteria

Studies were selected using the PICOS framework (see Table 1). Clinical populations were identified through diagnostic evidence, clinical settings, or appropriate screening measures. Studies were included if they provided diagnostic evidence or were conducted in a clinical setting, even if screening measures were absent. Seven papers were determined to be from a clinical population based on diagnostic evidence or setting. Key inclusion criteria were as follows:

- Adult clinical populations (≥ 18 years)
- CFT interventions following Gilbert's model
- Quantitative measures of compassion flows, self-criticism, or shame
- Any study design except single case studies
- Published and unpublished studies to reduce publication bias

Studies needed to be available in English for inclusion due to practical limitations in translation resources and to ensure accurate data extraction and quality assessment.

Table 1. Inclusion and exclusion criteria within the PICOS framework.

PICOS	Inclusion	Exclusion
Population	- Adults aged 18 and above. - Pre-existing mental health condition and/or meeting the criteria of a mental health condition in accordance with a diagnostic tool (for example, DSM-5, ICD-10) indicated by a senior clinician, psychiatrist, and/or medical records. - Recruited from clinical mental health settings (for example, services ranging from maximum-secure to inpatient and community settings). - Scoring above a cut-off point on a relevant screening measure (for example, the General Anxiety Disorder-7 measure for anxiety).	Non-clinical populations. This includes - Substance misuse disorders treated within specialist substance misuse services. - Studies without reported mental health diagnoses and with screening measure means below 1 standard deviation of clinical cut-offs.
Intervention	Group or individual CFT derived from Paul Gilbert (2009, 2014, 2020), including psychoeducation (for example, the tricky brain, three regulatory systems, fear of compassion, and the role of shame and self-criticism) and exercises (for example, soothing rhythm breathing, mindful attention, and compassionate letter writing and/or imagery).	Any other interventions that do not constitute as CFT or do not cover the key components of the model, including other compassion-based interventions (for example, Neff's mindful self-compassion, mindfulness-based cognitive therapy, and compassion cultivation training).
Comparison	- Any alternative psychological intervention or control group (e.g., waitlist control or treatment-as-usual). - Studies with no comparator.	
Outcome	Pre- and post-intervention outcomes for: - Self-compassion, compassion to and from others; - Self-criticism/self-reassurance; - Internal and/or external shame.	- Studies that do not measure CFT-based outcomes. - Studies that measure physiological outcomes only, such as heart rate variability.

Table 1. Cont.

PICOS	Inclusion	Exclusion
Study Design	• Cohort studies. • Non-randomised designs. • Controlled clinical trials. • Randomised controlled trials, feasibility studies, and pilot studies. • Single case experimental designs.	• Absence of pre–post statistical analysis or reliable change indices. • Case studies ($n = 1$). • Cross-sectional studies. • Studies using only qualitative methodologies.

2.2. Search Strategy

A comprehensive search strategy was developed using the PICOS framework (Methley et al., 2014) in consultation with a research librarian. The searches covered three major databases (PubMed, PsycInfo EBSCOhost, and Web of Science) from 2000 to 2024, using carefully selected terms for CFT interventions and outcomes of interest. A backward citation search was also conducted. Databases were searched in August 2023, February 2024, and October 2024. Search terms combined “Compassion focused therapy” variants with outcome-related terms (e.g., shame, self-criticism, and compassion flows). The complete search methodology, including all terms, Boolean operators, and database-specific adaptations, is provided in Supplementary Materials.

While the protocol indicated inclusion of unpublished dissertations and theses to reduce publication bias, the search was limited to three major academic databases (PubMed, PsycInfo EBSCOhost, Web of Science). Although PsycInfo EBSCOhost and Web of Science include selected unpublished dissertations, comprehensive grey literature searching would require additional specialised databases such as ProQuest Dissertations & Theses Global, which was beyond the scope of this review. All 21 studies included in the final synthesis were peer-reviewed published articles.

2.3. Selection Process

The study selection followed a systematic, multi-stage process to identify relevant CFT intervention research. Initial database searches of PubMed, PsycInfo EBSCOhost, and Web of Science identified 1209 records, supplemented by 1208 records from backward citation searching. After removing duplicates, 1724 unique records underwent title and abstract screening against eligibility criteria.

Full-text assessment of 66 potentially eligible reports led to 21 studies meeting the inclusion criteria. Primary reasons for exclusion included non-clinical populations ($n = 26$), incorrect outcomes ($n = 13$), and inappropriate interventions ($n = 1$). Five papers could not be retrieved for full-text review despite attempts to contact the authors, and these were excluded from the analysis. Author contact clarified the eligibility of three studies. All 21 studies included in the final synthesis underwent a complete full-text review.

To ensure selection reliability, an independent reviewer (a trainee clinical psychologist) screened 10% of the database records ($n = 66$), selected through random number generation. This process achieved complete agreement ($\kappa = 1.00$), supporting the robustness of the screening procedure.

A PRISMA flow diagram detailing the selection process and search strategy is provided in Figure 1.

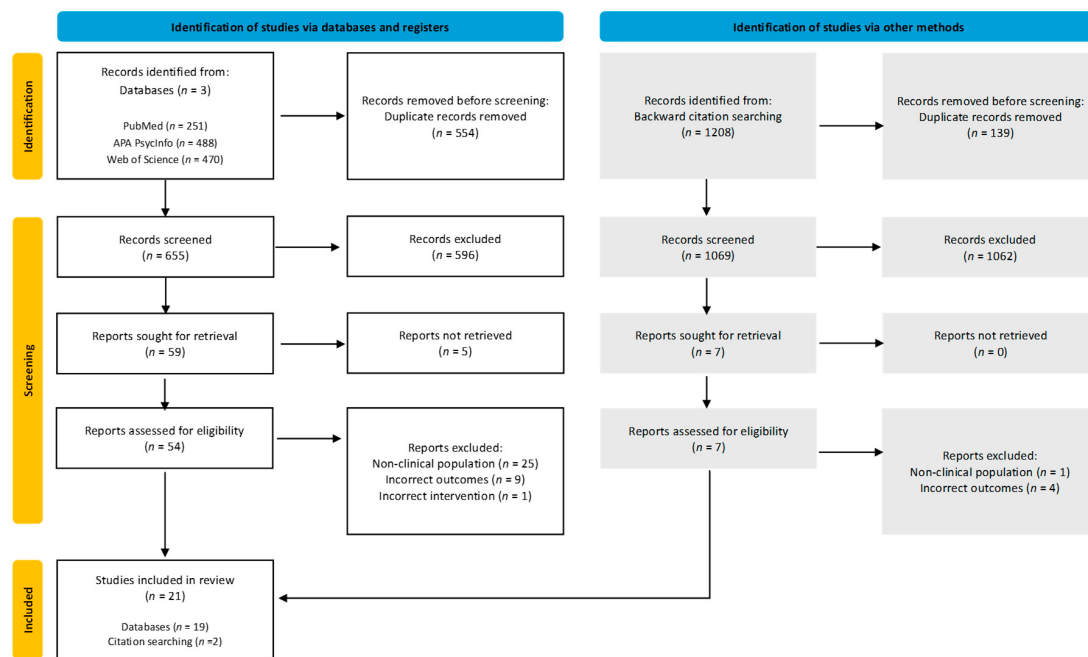


Figure 1. PRISMA flow diagram of the search strategy.

2.4. Data Extraction

Data extraction covered publication details, study characteristics, sample demographics, intervention details, study design, outcome measures, and outcomes. The studies were assessed for effectiveness using statistical significance and standardised mean differences at post-intervention and follow-up. All data were manually extracted and tabulated.

One study did not report participant gender data (Altavilla & Strudwick, 2022). Despite author contact, this information remained unavailable; however, the study was retained as it met the inclusion criteria and included a three-month follow-up assessment, which provided valuable longitudinal data for the synthesis.

Outcomes were categorised into three domains: compassion flows (self-compassion, compassion from others, and compassion to others), self-criticism (including inadequate self, hated self, and self-reassurance), and shame (internal and external).

Effect sizes were computed by the authors for all studies using available statistics (means, standard deviations, sample sizes, test statistics) to ensure precision and comparability across studies. Effect sizes were calculated using reported statistics following established methods (Beckham, 2016). Authors were contacted for missing data when necessary.

Effect sizes were calculated using Cohen's *d* for parametric tests and Rosenthal's *r* for non-parametric tests, with Hedges' *g* correction for small samples (Cohen, 1988; Rosenthal, 1994; Hedges, 1981; Hedges et al., 2012). Standard interpretations were used for effect size magnitude (very small, small, medium, large), as outlined in Table 2.

Table 2. Standardised mean difference classifications.

Effectiveness	Very Small	Small	Medium	Large
Cohen's <i>d</i> (Cohen, 1988)	<0.20	0.20	0.50	≥0.80
Hedges' <i>g</i> (Cohen, 1988)		0.20	0.50	≥0.80
Rosenthal's <i>r</i> (Rosenthal, 1994)		0.10	0.30	≥0.50

2.5. Quality Assessment

Study quality was evaluated using the Quality Assessment Tool for Quantitative Studies developed by the Effective Public Health Practice Project (EPHPP; Thomas et al.,

2004), which assesses selection bias, study design, confounders, blinding, data collection methods, attrition, intervention integrity, and analyses. Intervention integrity and analyses do not contribute to the global study quality rating. Global ratings were assigned as strong (low risk), moderate (some concerns), or weak (high risk). The EPHPP tool uses the term ‘blinding’; however, this paper uses the term ‘masking’ for inclusivity. Both terms refer to procedures that keep assessors and/or participants unaware of group allocation or study hypotheses to reduce bias. Due to ethical constraints in psychological intervention research, the overall grading and grading excluding ‘masking’ were reported to offer a more equitable assessment.

An independent and anonymous reviewer (a trainee clinical psychologist) screened 20% of the 21 included reports ($n = 4$) against the quality assessment tool to ensure inter-rater reliability. During the first discussion, a complete agreement was reached (100%; $\kappa = 1.00$).

2.6. Synthesis Methods

Due to heterogeneity in study designs and outcomes, a narrative synthesis approach was adopted. Results were organised by outcome domain (compassion flows, self-criticism, and shame) and synthesised using descriptive statistics and effect sizes. Hedges’ g effect sizes were summarised visually using a forest-style plot. The forest plot illustrates the direction and magnitude of effects across studies but does not reflect a formal meta-analytic synthesis. Only studies in which Hedges’ g could be computed were included in the forest plot to ensure consistency and interpretability of the effect size estimates. Studies without sufficient data to compute Hedges’ g could not be visualised in this format and were therefore excluded from the figure. Study quality ratings were integrated into the synthesis to evaluate the certainty of evidence.

3. Results

3.1. Study Characteristics

The 21 included studies were conducted in nine countries: Australia ($n = 1$), Colombia ($n = 1$), Denmark ($n = 1$), Iran ($n = 4$), Italy ($n = 1$), Japan ($n = 1$), the Netherlands ($n = 1$), the UK ($n = 9$), and the USA ($n = 2$) from 2006 to 2024. Sample sizes ranged from 7 (Clapton et al., 2018) to 91 participants (Fox et al., 2021). Of the 450 participants, 72.88% identified as women ($n = 328$), aged 18 to 89. One study did not report gender (Altavilla & Strudwick, 2022).

The studies included eight adult clinical populations: depression ($n = 4$), personality disorder ($n = 3$), bipolar disorder ($n = 2$), post-traumatic stress disorder (PTSD; $n = 1$), anxiety ($n = 1$), schizophrenia/psychosis ($n = 1$), obsessive-compulsive disorder (OCD; $n = 1$), and prolonged grief disorder ($n = 1$). Seven studies included transdiagnostic samples. The study settings included community ($n = 7$), specialist services ($n = 6$), secondary mental health ($n = 5$), and inpatient facilities ($n = 3$).

Most studies were cohort designs ($n = 12$; Altavilla & Strudwick, 2022; Clapton et al., 2018; Fox et al., 2021; Gilbert et al., 2022; Gilbert & Procter, 2006; Grodin et al., 2019; Judge et al., 2012; Laithwaite et al., 2009; Lucre & Corten, 2013; McLean et al., 2022; McManus et al., 2018; Naismith et al., 2021), with five RCTs (Asano et al., 2022; Behrouzian et al., 2024; Gharraee et al., 2018; Johannsen et al., 2022; Savari et al., 2021), two multiple baseline designs (Petrocchi et al., 2021; Pol et al., 2024), one case-controlled clinical trial with a waitlist control (Noorbala et al., 2013), and one cohort analytic study with treatment-as-usual as the control group (Stroud & Griffiths, 2021).

Thirteen studies (61.90%) included follow-up assessments ranging from one month to one year. Twenty studies reported retention rates averaging 77.18% (range: 24.39–100%). The study characteristics are presented in Table 3, and the study designs and assessment time points are detailed in Table 4.

Table 3. Characteristics of included studies.

First Author	Clinical Group		Sample Size		Sample Details		Intervention (Reference)	Comparator/ Control Group	Individual or Group Delivery	Session Number, Duration in Hours (h), and Frequency	Facilitator(s)
Location	(Diagnostic Tool)	Study Diagnostic Interview (Yes/No); Diagnosis Date; Diagnostic Evidence	Included Sample	Retention Rate (%)	Age (M, SD, Range) Gender (% Female, % Male)	Facilitator Training					
Altavilla and Strudwick (2022) UK	Depression, anxiety, bipolar disorder, PTSD, and psychosis, recruited from an NHS secondary mental health setting. (NR cut-offs on CORE-OM, DASS-21)	No; Unspecified; Medical records.	23	22/23 (95.65%)	NR, NR, 32–89 NR, NR	CFT (Paul Gilbert, 2009, 2010 ; Lee & James, 2012 ; Welford, 2012)	N/A	Group (Four cohorts with 6–10 people)	20, 2.5 h, weekly	Two clinical psychologists. Clinical work with CFT.	
Asano et al. (2022) Japan	Treatment-resistant depression, recruited from the community. (≥20 on the BDI)	Yes; Unspecified; The Japanese M.I.N.I	17 CFT: <i>n</i> = 10 TAU: <i>n</i> = 7	Total: 17/18 (94.44%) CFT: 9/10 (90.00%) TAU: 7/7 (100.00%)	CFT: 39.80, 11.22, 24–56 80.00%, 20.00% TAU: 40.00, 11.46, 26–55 85.00%, 15.00%	CFT (Judge et al., 2012)	TAU	Group	12, 1.5 h, weekly	One clinical psychologist, one industrial counsellor. 1/2 facilitators attended 3-day CFT workshop. Both CBT trained and supervised by PG.	

Table 3. Cont.

First Author	Clinical Group		Sample Size		Sample Details	Intervention (Reference)	Comparator/ Control Group	Individual or Group Delivery	Session Number, Duration in Hours (h), and Frequency	Facilitator(s)
Location	(Diagnostic Tool)	Study Diagnostic Interview (Yes/No); Diagnosis Date; Diagnostic Evidence	Included Sample	Retention Rate (%)	Age (M, SD, Range) Gender (% Female, % Male)					Facilitator Training
Behrouzian et al. (2024) Iran	Bipolar disorder, recruited from a psychiatric clinic.	Yes; Unspecified; Bipolar disorder diagnosed by an expert clinician using the SCID-I	26 CFT: <i>n</i> = 13 TAU: <i>n</i> = 13	Total: 26/31 (86.66%) CFT: 13/16 (81.25%) TAU: 13/15 (86.66%)	CFT: 29.15, 6.44, NR 84.62%, 15.38% TAU: 28.46, 5.86, NR 76.92%, 23.08%	CFT (Gilbert, 2010, 2014; Gharraee et al., 2018)	TAU	Individual	10, NR, weekly	NR NR
Clapton et al. (2018) UK	Depression and anxiety, recruited from an NHS Intellectual Disabilities service. (NR diagnostic tools)	No; Unspecified; Medical records.	7	6/7 85.71%	38.5, 15.6, NR 66.70%, 33.30%	CFT (Gilbert, 2010, 2014; Gilbert & Irons, 2004; Gilbert & Procter, 2006; Heriot-Maitland et al., 2014)	N/A	Group (Two cohorts with three people)	6, 1.5 h, weekly	Two clinical psychologists, one trainee clinical psychologist. Lead facilitators with three years of CFT experience and supervision by PG.

Table 3. Cont.

First Author	Clinical Group		Sample Size		Sample Details		Intervention (Reference)	Comparator/ Control Group	Individual or Group Delivery	Session Number, Duration in Hours (h), and Frequency	Facilitator(s)
Location	(Diagnostic Tool)	Study Diagnostic Interview (Yes/No); Diagnosis Date; Diagnostic Evidence	Included Sample	Retention Rate (%)	Age (<i>M</i> , <i>SD</i> , Range) Gender (% Female, % Male)	Facilitator Training					
Fox et al. (2021) USA	Distress at a clinical level, recruited from university counselling services. (≥64 on the OQ-45)	No; Unspecified; Unspecified.	91	45/91 49.45%	22.70, NR, 18–29 Among 75 participants who completed pre-treatment measures: 73.50%, 26.50%.	CFT (Paul Gilbert, 2009, 2017; Gilbert & Choden, 2013).	N/A	Group (Eight cohorts of 7–14 people) One individual consolidation session every three weeks	12, 2 h, weekly	Five clinical psychologists. 3/5 facilitators trained and supervised by PG.	
Gharraee et al. (2018) Iran	Social anxiety disorder, recruited from clinical settings. (LSAS pre-treatment mean 74.1 and 73.1 for CFT and waitlist, respectively)	Yes; 2017–2018; SCID-I/CV conducted by two assessors (psychiatrist and clinical psychologist).	34 CFT: <i>n</i> = 17 Waitlist: <i>n</i> = 17	Total: 32/34 94.12% CFT: 17/17 100.00% Waitlist: 15/17 88.24%	CFT: 22.7, 4.58, NR 47.10%, 52.90% Waitlist: 22.00, 4.39, NR 47.10%, 52.90%	CFT for social anxiety disorder (Boersma et al., 2015 , social anxiety protocol informed by Gilbert, 2010, 2014; Gilbert & Procter, 2006).	Waitlist	Individual	12, 1 h, weekly	One clinical psychologist. Protocol adherence was assessed by an independent and anonymous clinical psychologist for 20% of sessions.	

Table 3. Cont.

First Author	Clinical Group		Sample Size		Sample Details	Intervention (Reference)	Comparator/ Control Group	Individual or Group Delivery	Session Number, Duration in Hours (h), and Frequency	Facilitator(s)
Location	(Diagnostic Tool)	Study Diagnostic Interview (Yes/No); Diagnosis Date; Diagnostic Evidence	Included Sample	Retention Rate (%)	Age (M, SD, Range) Gender (% Female, % Male)					Facilitator Training
Gilbert et al. (2022) UK	Bipolar disorder, recruited from a specialist NHS Bipolar Service.	No; Unspecified; Medical records.	13	6/13 46.15%	NR, NR, 31–60 40.00%, 60.00%	CFT (Gilbert, 2001, 2010)	N/A	Group	25, NR, ~weekly	Two clinical psychologists. Trained and supervised by PG.
Gilbert and Procter (2006) UK	Personality disorder, recruited from an NHS Day Centre. (HADS anxiety and depression in clinical range)	No; Unspecified; Psychiatric diagnosis diagnosed a psychiatrist.	9	6/9 66.67%	45.20, 5.54, 39–51 66.60%, 33.30%	CFT (Gilbert & Irons, 2004)	N/A	Group	12, 2 h, weekly	Facilitated by PG and a cognitive therapist.
Grodin et al. (2019) USA	PTSD, recruited from a specialist outpatient clinic. (≥31 on PCL-5)	No; Unspecified; Psychiatric diagnosis.	22	16/22 72.73%	52.60, 12.90, NR 4.00%, 96.00%	CFT (Kolts, 2015)	N/A	Group	12, NR, NR	Four clinical psychologists (two facilitators in each group). Trained and supervised by RK.

Table 3. Cont.

First Author	Clinical Group		Sample Size		Sample Details		Intervention (Reference)	Comparator/ Control Group	Individual or Group Delivery	Session Number, Duration in Hours (h), and Frequency	Facilitator(s)
Location	(Diagnostic Tool)	Study Diagnostic Interview (Yes/No); Diagnosis Date; Diagnostic Evidence	Included Sample	Retention Rate (%)	Age (M, SD, Range) Gender (% Female, % Male)	Facilitator Training					
Johannsen et al. (2022) Denmark	Prolonged grief disorder, recruited from a bereavement study. (≥25 on the PG-13)	No; Unspecified; Diagnostic tool.	Total: n = 82 CFT: n = 42 Waitlist: n = 40	Total: 61/82 74.39% CFT: 26/42 61.90% Waitlist: 35/42 83.33%	Total: 60.50, 13.40, 23–83 67.10%, 32.90% CFT: 61.40, 13.60, NR 71.40%, 28.60% Waitlist: 59.50, 13.80, NR 62.5%, 37.5%	CFT adapted for prolonged grief. (Paul Gilbert, 2009, 2014)	Waitlist	Group	8, 2.25 h, weekly	Three clinical psychologists. Trained and supervised by PG.	
Judge et al. (2012) UK	Transdiagnostic, recruited from NHS Community Mental Health Teams. (≥17 on the BDI)	No; Unspecified; Medical records.	42	36/42 completed sessions: 85.71%. 27/42 completed measures: 64.29%.	40.90, 8.80, 22–56 59.30%, 40.70%	CFT (Paul Gilbert, 2009)	N/A	Group (Seven cohorts of ~five people)	12–14, 2.25 h, weekly	NR Trained and supervised by PG.	

Table 3. Cont.

First Author	Clinical Group		Sample Size		Sample Details		Intervention (Reference)	Comparator/ Control Group	Individual or Group Delivery	Session Number, Duration in Hours (h), and Frequency	Facilitator(s)
Location	(Diagnostic Tool)	Study Diagnostic Interview (Yes/No); Diagnosis Date; Diagnostic Evidence	Included Sample	Retention Rate (%)	Age (M, SD, Range) Gender (% Female, % Male)	Facilitator Training					
Laithwaite et al. (2009) UK	Schizophrenia, Schizo-affective or bipolar affective disorder, recruited from a maximum-secure hospital. (NR)	No; Unspecified; Unspecified.	19	18/19 94.74%	36.90, 9.10, NR 0.00%, 100.00%	CFT (Gilbert, 2001)	N/A	Group (Three cohorts)	20, NR, twice weekly	Two clinical psychologists, one advanced practitioner, one trainee clinical psychologist, and two assistant psychologists. NR	
Lucre and Corten (2013) UK	Personality disorder, recruited from a specialist NHS Personality Disorder Service.	Yes; 2010; International PD examination completed by an NHS Cognitive Behavioural Psychotherapist.	10	8/10 80.00%	NR, NR, 18–54 77.70%, 22.20%	CFT (Gilbert, 2001)	N/A	Group	16, NR, NR	Two psychotherapists. Trained and supervised by PG.	

Table 3. Cont.

First Author	Clinical Group		Sample Size		Sample Details					Facilitator(s)
Location	(Diagnostic Tool)	Study Diagnostic Interview (Yes/No); Diagnosis Date; Diagnostic Evidence	Included Sample	Retention Rate (%)	Age (<i>M, SD, Range</i>) Gender (% Female, % Male)	Intervention (Reference)	Comparator/ Control Group	Individual or Group Delivery	Session Number, Duration in Hours (h), and Frequency	Facilitator Training
McLean et al. (2022) Australia	Transdiagnostic, recruited from specialist Sexual Assault services. (≥ 31 on PCL-5 or ≥ 10 and ≥ 8 on DASS-21 on depression and anxiety subscales, respectively).	No; No; Unspecified.	36	Pre to post-CFT: 30/36 83.33% Pre to follow-up: 25/36 69.44%	41.10, 13.30, 18–65 100.00%, 0.00%	CFT (Lee & James, 2012 ; Paul Gilbert, 2009, 2015)	N/A	Group (Four cohorts of 7–11 people)	12, 2 h, weekly	One psychologist, one counsellor. Psychologist had 20 years’ clinical experience working with survivors of interpersonal trauma; trained and supervised by PG.
McManus et al. (2018) UK	Transdiagnostic, recruited from NHS Community Mental Health Teams. NR	No; Unspecified; Medical records.	27	13/27 48.15%	43, NR, 23–67 46.20%, 53.80%	CFT (Braehler et al., 2013 ; Gilbert, 2010 ; Welford, 2012)	N/A	Group (Four cohorts) One individual session (week 5)	16, 2 h, weekly	Two facilitators per group: either clinical psychologists or trainee clinical psychologists. NR

Table 3. Cont.

First Author	Clinical Group		Sample Size		Sample Details		Intervention (Reference)	Comparator/ Control Group	Individual or Group Delivery	Session Number, Duration in Hours (h), and Frequency	Facilitator(s)
Location	(Diagnostic Tool)	Study Diagnostic Interview (Yes/No); Diagnosis Date; Diagnostic Evidence	Included Sample	Retention Rate (%)	Age (<i>M, SD, Range</i>) Gender (% Female, % Male)	Facilitator Training					
Naismith et al. (2021) Colombia	Transdiagnostic, recruited from non-governmental organisations for low-income women. (either ≥ 34 on IES-R, or ≥ 10 on PHQ-9, or ≥ 8 on GAD-7).	No; Unspecified; Unspecified.	41	10/41 24.39%	37.80, 11.70, 19–52 100.00%, 0.00%	CFT (Gilbert, 2014)	N/A	Group	5, 2.25 h, weekly	One clinical psychologist, one marriage and family therapist. NR	
Noorbala et al. (2013) Iran	Depression, recruited from a psychiatric clinic. (≥ 20 on BDI)	Yes; Unspecified; Psychiatrist diagnosed major depressive disorder according to DSM-IV criteria.	Total: $n = 22$ CFT: $n = 11$ Waitlist: $n = 11$	Total: 19/22 86.36% CFT: 9/11 81.82% Waitlist: 10/11 90.91%	28.20, NR, 20–40 100.00%, 0.00%	CFT (Gilbert, 2005)	Waitlist	Group	12, 2 h, twice weekly	NR NR	

Table 3. Cont.

First Author	Clinical Group		Sample Size		Sample Details		Intervention (Reference)	Comparator/ Control Group	Individual or Group Delivery	Session Number, Duration in Hours (h), and Frequency	Facilitator(s)
Location	(Diagnostic Tool)	Study Diagnostic Interview (Yes/No); Diagnosis Date; Diagnostic Evidence	Included Sample	Retention Rate (%)	Age (M, SD, Range) Gender (% Female, % Male)	Facilitator Training					
Petrocchi et al. (2021) Italy	Obsessive compulsive disorder, recruited from an anxiety and mood disorders unit. (≥14 on Y-BOCS)	Yes; Unspecified; OCD diagnosed by an expert clinician using the SCID-I	8	8/8 100.00%	NR, NR, 34–41 50.00%, 50.00%	CFT (Gilbert & Choden, 2013)	N/A	Group	8, 2 h, weekly	Two psychotherapists. One with 8 years of training with PG, one with five years of training in CFT.	
Pol et al. (2024) The Netherlands	Personality disorder, recruited from a day-hospital. (≥14.93 on BPDSI-IV)	No; Unspecified; Unspecified	12	9/12 75.00%	39.30, 44.30, NR 100.00%, 0.00%	CFT (Gilbert et al., in press ; Pol et al., 2020)	N/A	Group	12, NR, weekly	One clinical psychologist. Basic and advanced training in CFT with PG.	

Table 3. Cont.

First Author	Clinical Group		Sample Size		Sample Details					Facilitator(s)
Location	(Diagnostic Tool)	Study Diagnostic Interview (Yes/No); Diagnosis Date; Diagnostic Evidence	Included Sample	Retention Rate (%)	Age (<i>M, SD, Range</i>) Gender (% Female, % Male)	Intervention (Reference)	Comparator/ Control Group	Individual or Group Delivery	Session Number, Duration in Hours (h), and Frequency	Facilitator Training
Savari et al. (2021) Iran	Major depressive disorder, recruited from university counselling services. (≥20 on BDI)	Yes; NR; SCID-I	Total: 30 CFT: 15 Waitlist: 15	Total: 30/30 100.00% CFT: 15/15 100.00% Waitlist: 15/15 100.00%	24.30, 2.16, 21–29 100.00%, 0.00%	CFT (Gilbert, 2010; Gilbert & Choden, 2013)	N/A	Group	8, 1.5 h, twice weekly for 4 weeks	Senior author. Attended CFT training with PG.
Stroud and Griffiths (2021) UK	Transdiagnostic, recruited from an inpatient mental health setting. (≥10 on CORE-OM)	No; Unspecified; Unspecified.	Total: <i>n</i> = 32 CFT: <i>n</i> = 19 TAU: <i>n</i> = 13	NR	NR, NR, 18–60 59.40%, 40.60%	TAU & CFT (Gilbert, year unspecified)	TAU	Group (open format)	6, 1 h, daily (Repeated cyclically every week over 4-months)	NR NR

Notes: NR = Not Reported; PG = Paul Gilbert; RK = Russel Kolts; CORE-OM = Clinical Outcomes in Routine Evaluation—Outcome Measure (Evans et al., 2000); BDI = Beck's Depression Inventory (Beck et al., 1993); Japanese M.I.N.I = Mini-international Neuropsychiatric Interview (Otsubo et al., 2005); DASS-21 = Depression Anxiety and Stress Scale (Lovibond & Lovibond, 1995); OQ-45 = Outcome Questionnaire-45 (Lambert et al., 2004); SCID-I/CV = Structured Clinical Interview for DSM-IV Axis I Disorders, Clinical Version (First et al., 1996); PG-13 = Prolonged Grief-13 (Prigerson et al., 2009); PCL-5 = PTSD Checklist for DSM-5 (Weathers et al., 2014); IES-R = Impact of Event Scale-Revised (Weiss & Marmar, 1997); PHQ-9 = Patient Health Questionnaire (Kroenke et al., 2001); GAD-7 = Generalised Anxiety Disorder Assessment (Spitzer et al., 2006); Y-BOCS = Yale-Brown Obsessive-Compulsive Scale (Goodman et al., 1989); BPDSI-IV = The Borderline Personality Disorder Severity Index-IV (Arntz et al., 2003); SCID-I = Structured Clinical Interview for Diagnosis (First & Gibbon, 2004).

Table 4. Results across studies.

First Author Location	Design	Outcome Measures (Measuring Tools)	Timepoints	Main Outcome(s) (Reported <i>p</i> , ES (<i>d</i>), and Calculated <i>g</i>)
Altavilla and Strudwick (2022) UK	Mixed methods (Cohort: One group pre + post, and qualitative interviews)	Self-compassion (SCS)	Pre-intervention, Session 20, 3-month follow-up	^a Significant increase in self-compassion (SCS) post-intervention (<i>p</i> = 0.008)
Asano et al. (2022) Japan	RCT (CFT vs. TAU)	Self-compassion (CEAS) Compassion from others (CEAS) Compassion to others (CEAS)	Pre-intervention, post-intervention	Change in self-compassion (CEAS) post-intervention (<i>d</i> = 1.08 ^L , <i>g</i> = 1.13 ^L) compared to TAU. Change in compassion from others (CEAS) post-intervention (<i>d</i> = 0.61 ^M , <i>g</i> = 0.66 ^M) compared to TAU. Change in compassion to others (CEAS) post-intervention (<i>d</i> = 0.14 ^{VS} , <i>g</i> = 0.15 ^{VS}) compared to TAU.
Behrouzian et al. (2024) Iran	RCT (CFT vs. TAU)	Self-compassion (SCS-SF) Self-criticism (SCRS)	Pre-intervention, post-intervention, and 2-month follow-up	No significant change in self-compassion (SCS-SF) post-intervention compared to TAU (<i>p</i> = 0.49, <i>d</i> = 0.50 ^M , <i>g</i> = 0.48 ^S). Significant decrease in self-criticism (SCRS) post-intervention compared to TAU (<i>p</i> < 0.001, <i>d</i> = −1.60 ^L , <i>g</i> = −1.55 ^L).
Clapton et al. (2018) UK	Mixed methods (Cohort: One group pre + post, and qualitative interviews)	Self-compassion (SCS-SF) Self-criticism (SCS-SF)	Pre-intervention, 2–4 weeks post-intervention	No significant change in self-compassion (SCS-SF) post-intervention (<i>p</i> = 0.674, <i>d</i> = 0.25 ^S , <i>g</i> = 0.23 ^S). Significant decrease in self-criticism (SCS-SF) post-intervention (<i>p</i> = 0.027, <i>d</i> = −2.04 ^L , <i>g</i> = −1.88 ^L).

Table 4. Cont.

First Author Location	Design	Outcome Measures (Measuring Tools)	Timepoints	Main Outcome(s) (Reported <i>p</i> , ES (<i>d</i>), and Calculated <i>g</i>)
Fox et al. (2021) USA	Cohort: One group pre + post	Self-compassion (CEAS) Compassion from others (CEAS) Compassion to others (CEAS) Self-criticism (Inadequate Self, Hated Self & Reassured Self; FSCRS) Self-criticism (DEQ)	Pre-intervention, session 6, post-intervention (session 12+)	Significant increase in self-compassion (CEAS) post-intervention ($p < 0.001$, $d = 0.75^M$, $g = 0.67^M$).
				Significant increase in compassion from others (CEAS) post-intervention ($p < 0.001$, $d = 0.26^S$, $g = 0.26^S$).
				No significant change in compassion to others (CEAS) post-intervention ($p = 0.19$, $d = -0.08^{VS}$, $g = -0.08^{VS}$).
				Significant decrease in inadequate self and hated self (FSCRS) post-intervention ($p < 0.001$, $d = -0.71^M$, $g = -0.70^M$, and $p < 0.001$, $d = -0.34^S$, $g = -0.33^S$, respectively).
				Significant increase in reassured self (FSCRS) post-intervention ($p < 0.001$, $d = 0.40^S$, $g = 0.40^S$).
Gharraee et al. (2018) Iran	RCT (CFT vs. Waitlist)	Self-compassion (SCS) Self-criticism (LOSC)	Pre-intervention, post-intervention, and 2-month follow-up	Significant decrease in self-criticism (DEQ) post-intervention ($p < 0.001$, $d = -0.50^M$, $g = -0.50^M$).
				Significant increase in self-compassion (SSC) compared to waitlist control at post-intervention ($p < 0.001$, $d = 1.99^L$, $g = 1.94^L$), which was maintained at follow-up ($p = 0.39$, $d = -0.37^S$, $g = -0.36^S$).
				Significant decrease in self-criticism (LOSC) compared to waitlist control at post-intervention ($p < 0.001$, $d = -1.91^L$, $g = -1.86^L$), which further increased at follow-up ($p < 0.001$, $d = 0.26^S$, $g = 0.25^S$).

Table 4. Cont.

First Author Location	Design	Outcome Measures (Measuring Tools)	Timepoints	Main Outcome(s) (Reported p , ES (d), and Calculated g)
Gilbert et al. (2022) UK	Mixed methods (Cohort: One group pre + post, and qualitative interviews)	Self-compassion (CEAS) Compassion from others (CEAS) Compassion to others (CEAS) Self-criticism (Inadequate Self, Hated Self & Reassured Self; FSCRS)	Pre-intervention, 12 weeks, 32 weeks, and post-intervention (45 weeks)	An overall improvement in self-compassion and compassion from others (CEAS) were reported from baseline to 45 weeks (6 out of 6 participants improved, $d = 1.29^L$, $g = 1.19^L$, and 4 out of 6 participants improved, $d = 0.39^S$, $g = 0.36^S$, respectively).
				No overall improvement in compassion to others (CEAS) was reported from baseline to 45 weeks (3 out of 6 participants improved, $d = -0.38^S$, $g = -0.35^S$).
				An overall improvement in inadequate self, hated self, and reassured self (FSCRS) were reported from baseline to 45 weeks (6 out of 6 participants improved, $d = -1.21^L$, $g = -1.11^L$, 2 out of 6 participants improved, $d = -0.21^S$, $g = -0.19^{VS}$, and 5 out of 6 improved, $d = 0.61^M$, $g = 0.56^M$, respectively).
Gilbert and Procter (2006) UK	Cohort: One group pre + post	Self-compassion (Weekly Interval Contingent Diary Measuring Self-attacking and Self-soothing) Self-criticism (Inadequate Self, Hated Self & Reassured Self; FSCRS) External shame (OAS)	Pre-intervention (beginning of week 1), post-intervention (end of week 12)	Significant increase in self-compassion (Diary) post-intervention ($p = 0.030$, $d = 4.49^L$, $g = 4.14^L$).
				No significant change in inadequate self (FSCRS) post-intervention ($p = 0.070$, $d = -2.73^L$, $g = -2.52^L$).
				Significant decrease in hated self (FSCRS) post-intervention ($p = 0.030$, $d = -2.04^L$, $g = -1.88^L$).
				Significant increase in reassured self (FSCRS) post-intervention ($p = 0.030$, $d = 1.86^L$, $g = 1.71^L$).
				Significant decrease in external shame (OAS) post-intervention ($p = 0.030$, $d = -0.82^L$, $g = -0.75^M$).

Table 4. Cont.

First Author Location	Design	Outcome Measures (Measuring Tools)	Timepoints	Main Outcome(s) (Reported p , ES (d), and Calculated g)
Grodin et al. (2019) USA	Cohort: One group pre + post (pilot study)	Self-compassion (SCS)	Pre- and post-intervention	No significant change in self-compassion (SCS) post-intervention ($p = 0.34$, $d = 0.17^{VS}$, $g = 0.11^{VS}$).
Johannsen et al. (2022) Denmark	RCT (CFT vs. Waitlist)	Self-criticism (Inadequate Self, Hated Self & Reassured Self; FSCRS)	Pre-intervention, post-intervention, 3-month and 6-month follow-up	No significant change in inadequate self and hated self (FSCRS) compared to waitlist post-intervention ($p = 0.24$, $d = -0.44^S$, $g = -0.43^S$ and $p = 0.15$, $d = -0.08^{VS}$, $g = -0.08^{VS}$, respectively). Significant increase in reassured self (FSCRS) compared to waitlist post-intervention ($p = 0.001$, $d = 0.43^S$, $g = 0.43^S$).
Judge et al. (2012) UK	Cohort: One group pre + post	Self-compassion (Weekly Interval Contingent Diary Measuring Self-attacking and Self-soothing) Self-criticism (Inadequate Self, Hated Self & Reassured Self; FSCRS) Internal shame (ISS) External shame (OAS)	Pre- and post-intervention	Significant increase in self-compassion (Diary) post-intervention ($p < 0.001$, $d = 1.17^L$, $g = 1.15^L$). Significant decrease in inadequate self and hated self (FSCRS) post-intervention ($p < 0.001$, $d = -1.35^L$, $g = -1.33^L$ and $p < 0.001$, $d = -0.90^L$, $g = -0.88^L$, respectively). Significant increase in reassured self (FSCRS) post-intervention ($p < 0.001$, $d = 0.93^L$, $g = 0.91^L$). Significant decrease in internal shame (ISS) post-intervention ($p < 0.001$, $d = -1.30^L$, $g = -1.28^L$). Significant decrease in external shame (OAS) post-intervention ($p < 0.001$, $d = -0.55^M$, $g = -0.54^M$).
Laithwaite et al. (2009) UK	Cohort: One group pre + post	Self-compassion (SCS) External shame (OAS)	Pre-intervention, mid-group (5 weeks), post-intervention, 6-week follow-up	^a No significant change in self-compassion (SCS) post-intervention ($p = 0.180$, $r = 0.22^S$). ^a Significant decrease in external shame (OAS) post-intervention ($p = 0.040$, $r = 0.04^{<S}$).

Table 4. Cont.

First Author Location	Design	Outcome Measures (Measuring Tools)	Timepoints	Main Outcome(s) (Reported <i>p</i> , ES (<i>d</i>), and Calculated <i>g</i>)
Lucre and Corten (2013) UK	Mixed methods (Cohort: One group pre + post, and qualitative interviews)	Self-criticism (Inadequate Self, Hated Self & Reassured Self; FSCRS) External shame (OAS)	Pre- and post-intervention (16 weeks), 1-year follow-up	^a No significant change in inadequate self (FSCRS) at follow-up ($p = 0.062$).
				^a Significant decrease in hated self (FSCRS) post-intervention, which was maintained at follow-up ($p < 0.001$).
				^a Significant increase in reassured self (FSCRS) post-intervention, which was maintained at follow-up ($p < 0.001$).
				^a Significant decrease in external shame (OAS) post-intervention, which further decreased at follow-up ($p = 0.011$).
McLean et al. (2022) Australia	Cohort: One group pre + post	Self-compassion (CEAS) Self-criticism (Inadequate Self, Hated Self & Reassured Self; FSCRS) Global shame (EISS) External shame (OAS)	Pre-intervention (2 weeks pre-group), post-intervention (12 weeks), 3-month follow-up	Significant increase in self-compassion (CEAS) post-intervention ($p < 0.001$, $d = 1.82^L$, $g = 1.79^L$) and at follow-up ($p < 0.001$, $d = 1.57^L$, $g = 1.55^L$).
				Significant decrease in inadequate self (FSCRS) post-intervention ($p < 0.001$, $d = -1.59^L$, $g = -1.56^L$) and at follow-up ($p < 0.001$, $d = -1.73^L$, $g = -1.70^L$).
				Significant decrease in hated self (FSCRS) post-intervention ($p < 0.001$, $d = -0.98^L$, $g = -0.97^L$) and at follow-up ($p < 0.001$, $d = -0.94^L$, $g = -0.92^L$).
				Significant increase in reassured self (FSCRS) post-intervention ($p < 0.001$, $d = 1.20^L$, $g = 1.18^L$) and at follow-up ($p < 0.001$, $d = 1.11^L$, $g = 1.09^L$).
				Significant decrease in global shame (EISS) post-intervention ($p < 0.001$, $d = -1.70^L$, $g = -1.67^L$) at follow-up ($p < 0.001$, $d = -1.35^L$, $g = -1.32^L$).

Table 4. Cont.

First Author Location	Design	Outcome Measures (Measuring Tools)	Timepoints	Main Outcome(s) (Reported p , ES (d), and Calculated g)
McLean et al. (2022) Australia	Cohort: One group pre + post	Self-compassion (CEAS) Self-criticism (Inadequate Self, Hated Self & Reassured Self; FSCRS) Global shame (EISS) External shame (OAS)	Pre-intervention (2 weeks pre-group), post-intervention (12 weeks), 3-month follow-up	Significant decrease in external shame (OAS) post-intervention ($p < 0.001$, $d = -1.24^L$, $g = -1.22^L$) and at follow-up ($p < 0.001$, $d = -0.77^M$, $g = -0.75^M$).
McManus et al. (2018) UK	Mixed methods (Cohort: One group pre + post, and qualitative interviews)	Self-compassion (SCS) Self-criticism (Inadequate Self, Hated Self & Reassured Self; FSCRS) External shame (OAS)	Pre- and post-intervention (week 16)	^a Significant increase in self-compassion (SCS) post-intervention ($p = 0.010$). ^a Significant decrease in inadequate self (FSCRS) post-intervention ($p = 0.030$). ^a Significant decrease in hated self (FSCRS) post-intervention ($p = 0.020$). ^a No significant change in reassured self (FSCRS) post-intervention ($p = 0.070$). ^a Significant decrease in external shame (OAS) post-intervention ($p = 0.010$).
Naismith et al. (2021) Colombia	Cohort: One group pre + post	Self-criticism (Inadequate Self, Hated Self; FSCRS)	Pre- (5 weeks before intervention) and post-intervention, and 3-month follow-up	^b Inadequate self (FSCRS) reliably improved for 5 out of 10 participants post-intervention, and 5 out of 9 participants at follow-up. ^b Hated self (FSCRS) reliably improved for 4 out of 5 participants, and 3 out of 5 participants at follow-up (5 scored under 6 pre-intervention and were excluded).

Table 4. Cont.

First Author Location	Design	Outcome Measures (Measuring Tools)	Timepoints	Main Outcome(s) (Reported p , ES (d), and Calculated g)
Noorbala et al. (2013) Iran	Case controlled clinical trial (CFT vs. Waitlist)	Self-criticism (Comparative and Internalised; LOSC)	Pre-intervention, post-intervention, and 2-month follow-up	^a No significant change in comparative self-criticism (LOSC) post-intervention ($p = 0.491$) or at follow-up ($p = 0.108$). ^a No significant change in internalised self-criticism (LOSC) post-intervention ($p = 0.491$) or at follow-up ($p = 0.272$).
Petrocchi et al. (2021) Italy	Multiple baseline design	Self-compassion (Common Humanity; SCS) Self-criticism (Inadequate Self & Reassured Self; FSCRS)	Baseline, pre-intervention, post-intervention, and 1-month follow-up	No significant change in self-compassion (Common Humanity; SCS) post-intervention ($p = 0.11$, $g = 0.49^S$) and at follow-up ($p = 0.230$, $g = 0.57^M$). Significant decrease in inadequate self (FSCRS) post-intervention ($p = 0.030$, $g = -0.29^S$), which was not maintained at follow-up ($p = 0.110$, $g = -0.63^M$). Significant increase in reassured-self (FSCRS) post-intervention ($p = 0.040$, $g = 0.35^S$), which was not maintained at follow-up ($p = 0.100$, $g = 0.32^S$).
Pol et al. (2024) The Netherlands	Multiple baseline design	Self-compassion (SCS-SF) Self-criticism (FSCRS)	Baseline, pre-intervention, post-intervention, and 6-week follow-up	^b Self-compassion (SCS-SF) reliably improved for 3 out of 9 participants post-intervention and 5 out of 9 at follow-up. 5 out of 9 participants and 4 out of 9 participants showed no change post-intervention and at follow-up, respectively. ^b Self-criticism (FSCRS) reliably improved for 6 out of 9 participants post-intervention and at follow-up. 3 out of 9 participants showed no change.

Table 4. Cont.

First Author Location	Design	Outcome Measures (Measuring Tools)	Timepoints	Main Outcome(s) (Reported p , ES (d), and Calculated g)
Savari et al. (2021) Iran	RCT (CFT vs. Waitlist)	Self-compassion (SCS-SF) Self-criticism (FSCRS); Inadequate Self, Hated Self, & Reassured Self)	Pre- and post-intervention	Significant increase in self-compassion (SCS-SF) post-intervention compared to waitlist ($p = 0.010$, $d = 0.75^M$, $g = 0.73^M$).
				No significant change in inadequate self (FSCRS) post-intervention compared to waitlist.
				Significant decrease in hated self (FSCRS) post-intervention compared to waitlist ($p = 0.030$, $d = -1.00^L$, $g = -0.97^L$).
Stroud and Griffiths (2021) UK	Cohort analytic (Two group: CFT vs. TAU pre + post)	Self-compassion (VAS) Compassion to Others (VAS)	Pre-, every session, and post-intervention	Significant increase in reassured self (FSCRS) post-intervention compared to waitlist ($p < 0.001$, $d = 1.20^L$, $g = 1.17^L$).
				^a Significant increase in self-compassion (VAS) across all sessions compared to TAU ($p < 0.001$, $d = 0.24^S$).
				^a Significant increase in compassion to others (VAS) across all sessions compared to TAU ($p < 0.001$, $d = 0.18^{VS}$).

Notes: ^a d and/or g not calculated as information not available in full-text paper and no response from authors upon request for Supplementary Materials. ^b d and g are not calculated as the study describes the pattern of change or uses reliable change indices only. ^{VS}, ^S, ^M, and ^L indicate very small, small, medium, and large effect sizes, respectively. ^{<S} for Rosenthal's r indicates effect sizes under the recommended 0.10 cut-off for a small effect. SCS = The Self-Compassion Scale; SCS-SF = The Self-Compassion Scale Short Form; CEAS = Compassionate Engagement and Action Scales; SCRS = The Self-Critical Rumination Scale; FSCRS = Forms of Self-Criticism and Self-Reassuring Scale—Inadequate Self, Hated Self, and Reassured Self Subscales; DEQ = Depressive Experiences Questionnaire 48 McGill Revision—Self-Criticism Subscale; LOSC = Level of Self-Criticism Scale; OAS = the Other as Shamer Scale; ISS = Internalised Shame Scale; EISS = External and Internal Shame Scale; VAS = Visual Analogue Scale (0–10 Likert scale for self-compassion and compassion to others developed by authors).

3.2. Intervention Characteristics

All interventions followed Gilbert's CFT model (Gilbert, 2001, 2010; Gilbert & Choden, 2013; Gilbert & Irons, 2004; Gilbert & Procter, 2006; Gilbert et al., in press). Eight studies adapted protocols for specific populations (Clapton et al., 2018; Gharraee et al., 2018; Gilbert et al., 2022; Johannsen et al., 2022; Laithwaite et al., 2009; Lucre & Corten, 2013; Noorbala et al., 2013; Petrocchi et al., 2021).

Group therapy was predominant ($n = 19$), averaging 12.42 weekly sessions (range = 5–25), each lasting 1–2.5 h. Two studies used individual formats (Behrouzian et al., 2024; Gharraee et al., 2018). Qualified clinical psychologists delivered most interventions ($n = 14$), often with co-facilitators. Thirteen studies reported facilitator CFT training, with ten receiving ongoing supervision from Paul Gilbert (Asano et al., 2022; Clapton et al., 2018; Fox et al., 2021; Gilbert et al., 2022; Gilbert & Procter, 2006; Johannsen et al., 2022; Judge et al., 2012; Lucre & Corten, 2013; McLean et al., 2022; Petrocchi et al., 2021).

3.3. Quality Appraisal

3.3.1. Risk of Bias Within Studies

Initial quality assessment rated all studies as weak due to unavoidable participant awareness of intervention (the masking domain). Excluding masking, 14 studies were upgraded to moderate quality (see Figure 2; Behrouzian et al., 2024; Clapton et al., 2018; Fox et al., 2021; Gharraee et al., 2018; Gilbert et al., 2022; Gilbert & Procter, 2006; Johannsen et al., 2022; Judge et al., 2012; Laithwaite et al., 2009; Lucre & Corten, 2013; McLean et al., 2022; Noorbala et al., 2013; Petrocchi et al., 2021; Savari et al., 2021). Selection bias and confounders were consistently rated as weak across all studies. Six studies reported intervention consistency measures (Asano et al., 2022; Behrouzian et al., 2024; Fox et al., 2021; Gharraee et al., 2018; McLean et al., 2022; Savari et al., 2021). Three indicated a high likelihood of co-intervention, as participants received CFT alongside another intervention (Fox et al., 2021; Altavilla & Strudwick, 2022; Savari et al., 2021).

3.3.2. Risk of Bias Across Studies

Analysis across studies revealed consistent patterns in quality ratings. While confounders, data collection methods, withdrawals, and dropouts showed strong ratings, selection bias and masking were consistently rated as weak across all studies (see Figure 3). However, the study designs received a mostly moderate rating, indicating some bias concerns. This was due to recruitment methods and the inability to mask participants from psychological interventions. Reporting completeness varied, with several studies requiring author contact for additional methodological details.

3.4. Results Overview

Table 3 provides an overview of the study designs, outcome measures, assessment time points, and main findings for each included study. Effect sizes were computed for each study using standardised methods to maximise precision and ensure comparability. A forest-style summary of the standardised effect sizes (Hedges' g) is presented in Figure 4. In Figure 4, outcomes are grouped by construct, illustrating the magnitude and variability of CFT effects across compassion, self-criticism, and shame.

		Key Weak Moderate Strong Y Yes N No U Unclear										
Bias Risk	Study	Selection Bias	Study Design	Confounders	Blinding	Data Collection	Withdrawal and Dropout	Intervention Consistency Measured?	Co-intervention Likely?	Appropriate Analysis?	Global Rating	Global Rating without Blinding
Some Concerns	Behrouzian et al. (2024)							Y	U	Y		
	Clapton et al. (2018)							N	U	Y		
	Fox et al. (2021)							Y	Y	Y		
	Gharraee et al. (2018)							Y	N	Y		
	Gilbert et al. (2022)							N	N	Y		
	Gilbert and Procter (2006)							N	N	Y		
	Johannsen et al. (2022)							N	N	Y		
	Judge et al. (2012)							N	U	Y		
	Laithwaite et al. (2009)							N	U	Y		
	Lucre and Corten (2013)							N	N	Y		
	McLean et al. (2022)							Y	U	Y		
	Noorbala et al. (2013)							N	U	Y		
	Petrocchi et al. (2021)							N	U	Y		
	Savari et al. (2021)							Y	Y	Y		
High Risk of Bias	Altavilla and Strudwick (2022)							N	Y	Y		
	Asano et al. (2022)							Y	N	Y		
	Grodin et al. (2019)							U	N	Y		
	McManus et al. (2018)							N	U	Y		
	Naismith et al. (2021)							N	U	Y		
	Pol et al. (2024)							N	N	Y		
	Stroud and Griffiths (2021)							U	N	Y		

Figure 2. Quality appraisal within studies using the EPHPP tool. The figure includes studies by Altavilla and Strudwick (2022), Asano et al. (2022), Behrouzian et al. (2024), Clapton et al. (2018), Fox et al. (2021), Gharraee et al. (2018), Gilbert et al. (2022), Gilbert and Procter (2006), Grodin et al. (2019), Johannsen et al. (2022), Judge et al. (2012), Laithwaite et al. (2009), Lucre and Corten (2013), McLean et al. (2022), McManus et al. (2018), Naismith et al. (2021), Noorbala et al. (2013), Petrocchi et al. (2021), Pol et al. (2024), Savari et al. (2021), and Stroud and Griffiths (2021).

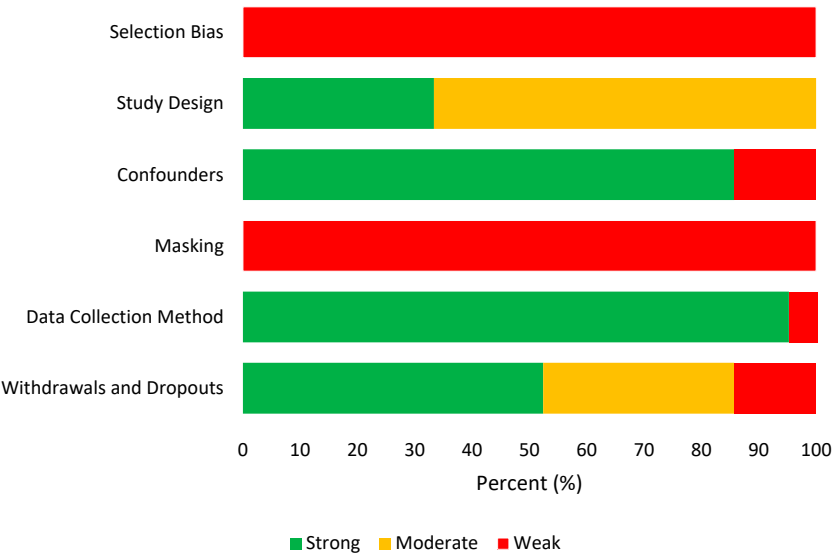


Figure 3. Quality appraisal across studies using the EPHPP tool.

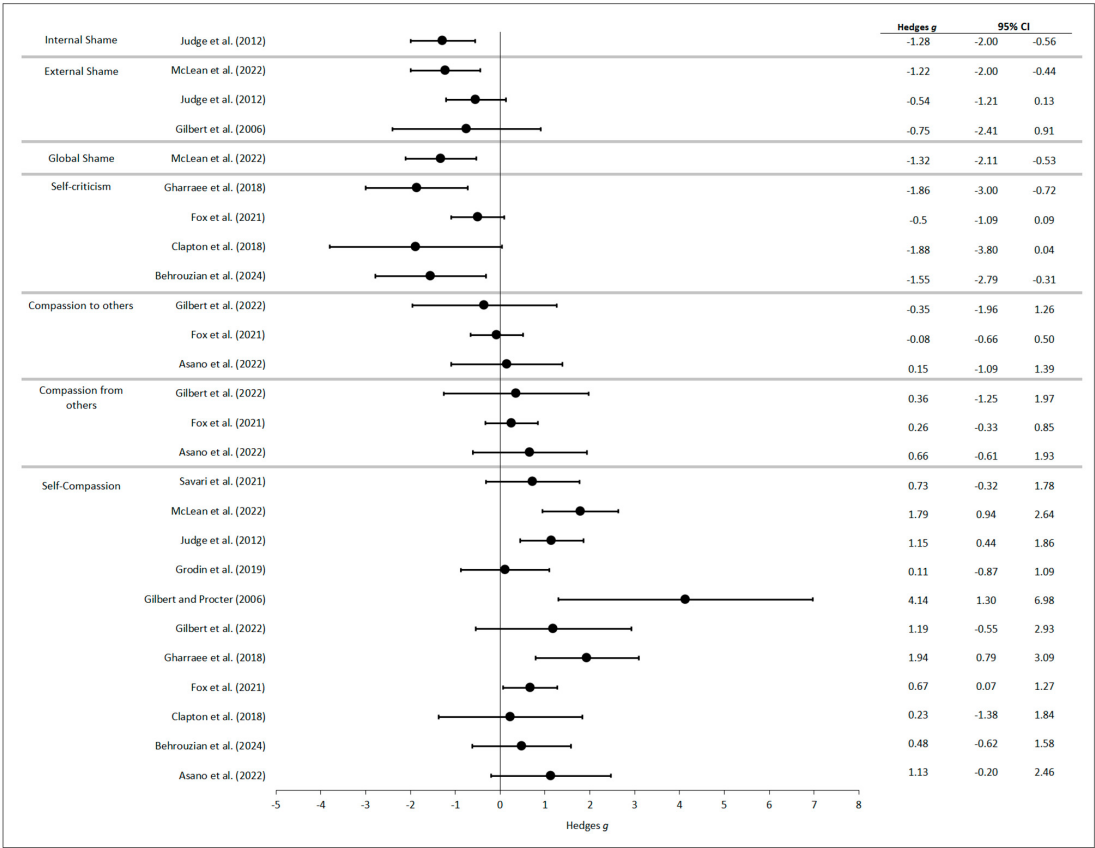


Figure 4. Forest plot summarising standardised effect sizes (Hedges' *g*) for CFT outcomes across the included studies. The figure includes studies by Asano et al. (2022), Behrouzian et al. (2024), Clapton et al. (2018), Fox et al. (2021), Gharraee et al. (2018), Gilbert et al. (2022), Gilbert and Procter (2006), Grodin et al. (2019), Judge et al. (2012), McLean et al. (2022), and Savari et al. (2021). Notes: Effect sizes (Hedges' *g*) are grouped by outcome: self-compassion, compassion from others, compassion to others, self-criticism, global shame, external shame, and internal shame. Error bars represent 95% confidence intervals. Positive values indicate an increase. Negative values indicate a decrease. Values reflect pre- to post-intervention changes. Only studies for which Hedges' *g* could be calculated based on available data are included. This figure provides a descriptive visual summary and is not based on meta-analysis.

3.4.1. Primary Outcomes

Compassion Flows

Measurement and Evidence Quality. Sixteen studies measured at least one compassion flow. Self-compassion was primarily assessed using validated measures: the Self-Compassion Scale (SCS; [Neff, 2003](#)), its short form (SCS-SF; [Raes et al., 2011](#)), and the Compassionate Engagement and Action Scales (CEAS; [Gilbert et al., 2017](#)). Some studies employed non-standardised measures like interval contingent diaries and visual analogue scales ([Gilbert & Procter, 2006](#); [Judge et al., 2012](#); [Stroud & Griffiths, 2021](#)). Study quality varied, with six moderate-quality and four weak-quality studies contributing to the self-compassion evidence base. The remaining compassion flows were measured less frequently (compassion from others: $n = 3$, compassion to others: $n = 4$), using either the CEAS or visual analogue scales, with studies ranging from weak to moderate quality ([Asano et al., 2022](#); [Fox et al., 2021](#); [Gilbert et al., 2022](#); [Stroud & Griffiths, 2021](#)). The underrepresentation of compassion to and from others reflects the study designs of the included research. The researchers predominantly selected self-compassion measures over available interpersonal compassion scales.

Effectiveness Results. Self-compassion showed the strongest evidence among compassion flows. Ten studies demonstrated significant improvements ($g = 0.23$ – 4.14 ; [Altavilla & Strudwick, 2022](#); [Asano et al., 2022](#); [Fox et al., 2021](#); [Gharraee et al., 2018](#); [Gilbert & Procter, 2006](#); [Judge et al., 2012](#); [McLean et al., 2022](#); [McManus et al., 2018](#); [Savari et al., 2021](#); [Stroud & Griffiths, 2021](#)). Four controlled studies showed superiority to waitlist or treatment-as-usual conditions ([Asano et al., 2022](#); [Gharraee et al., 2018](#); [Savari et al., 2021](#); [Stroud & Griffiths, 2021](#)). Follow-up data from four studies indicated maintenance of gains in most cases ([Gharraee et al., 2018](#); [McLean et al., 2022](#); [Petrocchi et al., 2021](#); [Pol et al., 2024](#)).

Evidence for the remaining compassion flows was more limited in both quantity and quality. Three studies examining the ability to receive compassion from others found small improvements ($g = 0.26$ – 0.36 ; [Asano et al., 2022](#); [Fox et al., 2021](#); [Gilbert et al., 2022](#)), with moderate-quality studies suggesting some reliability in these findings. However, the results for compassion to others were less convincing, with only two of four studies reporting very small positive effects ($d = 0.14$ – 0.18 ; [Asano et al., 2022](#); [Stroud & Griffiths, 2021](#)), and these came from methodologically weaker studies. This pattern of limited measurement, weaker methodology, and inconsistent effects prevents firm conclusions about CFT's impact on all compassion flows.

Self-Criticism

Measurement and Evidence Quality. Fifteen studies assessed self-criticism, primarily using the Forms of Self-Criticism/Self-Reassuring Scale (FSCRS; [Gilbert et al., 2004](#)). The FSCRS assesses self-criticism across three subscales: 'inadequate self', 'hated self', and 'reassured self'. Additional measures included the Level of Self-Criticism Scale (LOSC; [Thompson & Zuroff, 2004](#)) and the Depressive Experiences Questionnaire (DEQ; [Santor et al., 1997](#)). Most studies ($n = 9$) achieved moderate-quality ratings despite selection bias limitations ([Fox et al., 2021](#); [Gharraee et al., 2018](#); [Gilbert et al., 2022](#); [Gilbert & Procter, 2006](#); [Johannsen et al., 2022](#); [Judge et al., 2012](#); [Lucre & Corten, 2013](#); [McLean et al., 2022](#); [Petrocchi et al., 2021](#)). As such, there are some concerns about bias across self-criticism outcomes.

Effectiveness Results. Evidence for improvements in self-criticism was consistent across measures and studies. Of the eleven studies using the FSCRS, five reported significant reductions in 'inadequate self' ($g = 0.29$ – 1.56 ; [Fox et al., 2021](#); [Judge et al., 2012](#); [McLean et al., 2022](#); [McManus et al., 2018](#); [Petrocchi et al., 2021](#)), six showed decreased

'hated self' ($g = 0.33\text{--}0.97$; Fox et al., 2021; Gilbert & Procter, 2006; Judge et al., 2012; Lucre & Corten, 2013; McLean et al., 2022; McManus et al., 2018), and six demonstrated increased 'reassured self' ($g = 0.40\text{--}1.18$; Fox et al., 2021; Gilbert & Procter, 2006; Judge et al., 2012; Lucre & Corten, 2013; McLean et al., 2022; Petrocchi et al., 2021). Controlled studies from two studies somewhat strengthened these findings. Some of the subscales showed superiority to waitlist and treatment-as-usual conditions (Johannsen et al., 2022; Savari et al., 2021). Follow-up data from two studies showed mixed sustainability. One study reported large improvements at follow-up ($g = 0.92\text{--}1.70$; McLean et al., 2022), and another reported diminished effects (Petrocchi et al., 2021). A significant decrease in self-criticism was also found post-intervention in one study using the DEQ (Fox et al., 2021). Out of the two studies using the LOSC, only one found significant reductions compared to a waitlist control; however, the results were not maintained at follow-up (Gharraee et al., 2018). The improvements demonstrated across the studies suggest that CFT may reduce self-criticism in the short term; however, the moderate study quality indicates that results should be interpreted with caution.

Shame

Measurement and Evidence Quality. Seven studies examined shame outcomes, with most achieving moderate-quality ratings (Gilbert & Procter, 2006; Judge et al., 2012; Laithwaite et al., 2009; Lucre & Corten, 2013; McLean et al., 2022; McManus et al., 2018). External shame was assessed in six studies using the Other as Shamer Scale (OAS; Goss et al., 1994). Internal shame was measured in one study (Judge et al., 2012) using the Internalised Shame Scale (ISS; Cook, 1996). Global shame was measured in one study (McLean et al., 2022) using the External and Internal Shame Scale (EISS; Ferreira et al., 2022). As with the interpersonal compassion flows, internal shame was underrepresented in the included studies. The researchers typically selected external shame measures over tools assessing multiple dimensions of shame, such as the EISS (Ferreira et al., 2022). Even when the EISS was used (McLean et al., 2022), the available subscale data were not analysed. The predominance of moderate-quality studies suggests that most shame outcomes include some concerns with bias.

Effectiveness Results. External shame showed the most consistent evidence, with significant reductions across all six studies ($g = 0.54\text{--}1.22$; Gilbert & Procter, 2006; Judge et al., 2012; Laithwaite et al., 2009; Lucre & Corten, 2013; McLean et al., 2022; McManus et al., 2018). Five of these studies were of moderate quality, suggesting some concerns about bias and reduced confidence in these outcomes. Follow-up data from two studies showed either maintained (McLean et al., 2022) or enhanced reductions over time (Lucre & Corten, 2013). The single moderate-quality study measuring internal shame reported large reductions ($g = 1.28$; Judge et al., 2012). Global shame showed similarly large, sustained improvements ($g = 1.32\text{--}1.67$) in the one study measuring this construct (McLean et al., 2022). The limited measurement of these shame types prevents broader conclusions about CFT's effectiveness beyond external shame.

4. Discussion

This systematic review assessed the effectiveness of CFT interventions on the three flows of compassion, self-criticism, and shame in clinical populations. Across the 21 studies included, significant improvements in self-compassion were reported, with most studies showing medium to large effect sizes. Consistent reductions in self-criticism were also observed, particularly feelings of inadequacy and self-hatred. External shame decreased significantly in studies measuring this outcome, though internal shame was assessed less frequently. Evidence regarding compassion to others and receiving compassion from

others was limited and mixed. Some studies indicated maintained or increased effects at follow-up; however, few included comparator groups. Where comparisons were available, results suggested potentially greater improvements in self-compassion relative to waitlist and treatment-as-usual controls. By examining CFT interventions in relation to their core theoretical constructs, these findings address gaps in previous reviews that grouped a range of compassion-based interventions or focused solely on general symptom reduction.

4.1. Primary Findings

The review identified outcome patterns that partially support CFT's theoretical framework. Self-compassion and self-criticism emerged as the most reliably supported outcomes, likely reflecting their frequent inclusion across studies. The findings were relatively convergent across studies, though methodological limitations were common.

Self-compassion had the most comprehensive evidence base. Seventeen studies reported significant improvements, with effect sizes ranging from small to large ($g = 0.23$ to 4.14). Although most studies were rated as weak to moderate quality due to small samples, a high risk of bias, and limited demographic diversity, the consistency of findings increases confidence in this outcome. These results align with prior reviews highlighting CFT's efficacy in enhancing self-compassion (Kirby et al., 2017; Leaviss & Uttley, 2015). The findings support Gilbert's (2014) theory that self-compassion activates the soothing system, counteracting threat-based responses. This consistency likely reflects self-compassion's central role in CFT protocols and the use of validated measures. However, most studies ($n = 11$) used Neff's (2003) Self-Compassion Scale. Neff's (2003) scale focuses on self-directed compassion. Therefore, studies using this scale to measure self-compassion potentially overlook CFT's broader theoretical scope. The limited demographic diversity further restricts generalisability, with 72.88% of participants identifying as female across the nine countries represented. Overall, self-compassion is the most well-supported outcome in the current evidence base.

Self-criticism also showed consistent improvement across 14 studies, with effect sizes ranging from small to large ($g = 0.29$ to 1.56) across subscales, including inadequate self, hated self, and reassured self. These studies were typically of moderate quality, though common issues included small sample sizes and risk of bias. Despite this, the consistency supports CFT's effectiveness in reducing self-critical tendencies, echoing previous reviews (Leaviss & Uttley, 2015; Millard et al., 2023). These results align with theoretical claims that compassion practices reduce self-criticism by encouraging more balanced self-evaluation (Gilbert & Simos, 2022). However, the use of varied measures and subscales complicates direct comparisons. The gender imbalance noted for self-compassion also applies here. This limitation intersects with other methodological concerns such as selection bias, high attrition, and lack of masking (Craig et al., 2020; Millard et al., 2023). Although some evidence suggests gender differences in experiences of self-criticism (Rogier et al., 2023; Vanderheiden & Mayer, 2024), few studies have explored gender-based variation in CFT outcomes. Future research should investigate how people of different genders experience and respond to CFT to ensure that findings are relevant across diverse populations.

In contrast to the stronger findings for self-compassion and self-criticism, external shame was assessed in only six studies. These reported significant reductions, with medium to large effect sizes ($g = 0.54$ to 1.28). However, the studies were small and of moderate quality. While the results suggest that compassion may reduce shame, the mechanisms—such as increased social safeness or reduced fear of negative evaluation—were not directly examined (Gilbert, 2020). The limited number and quality of studies reduce confidence in this outcome compared to others. Demographic homogeneity further limits generalisability,

possibly obscuring cultural differences in the experience of shame. Thus, while the findings for external shame are promising, they remain less conclusive.

The evidence base is considerably weaker for the interpersonal components of Gilbert's model. Compassion to others and compassion from others were measured in only three to four studies each, with small and inconsistent effects. These studies had very small samples and were rated as weak to moderate quality. Several factors likely contribute to this limited focus: the relatively recent development of validated interpersonal compassion measures, practical constraints such as questionnaire length, the prioritisation of broad clinical outcomes over specific mechanisms, and the historical emphasis on self-directed compassion in CFT research. This lack of evidence presents challenges for clinicians, who currently have limited guidance on how CFT targets these relational aspects. Consequently, clinical choices about whether CFT should be delivered individually or as a group, due to the interpersonal nature of compassion, may be hindered. Gaining further clarity on interpersonal compassion will support the development of tailored approaches for clients in whom interpersonal dynamics are central to distress and recovery.

Internal shame has also received limited research attention. Only a handful of studies have examined this construct, typically with small samples and weak methodological designs. Although earlier reviews called for further investigation into CFT's impact on shame (Craig et al., 2020), internal shame remains underexplored. Even when subscale data were available, many studies reported only global shame scores. Additionally, concerns about participant burden and the acceptability of assessing shame, given its distressing and potentially shame-inducing nature, may deter inclusion in clinical research. This narrow focus limits clinicians' ability to understand how CFT influences internal shame. Such uncertainty hinders optimisation of treatment approaches, particularly for populations in whom internal shame plays a central role in maintaining psychological distress. Without clearer evidence, it is also difficult to determine whether CFT affects internal and external shame through similar or distinct mechanisms.

4.2. Appraisal of Included Evidence

Several methodological limitations affect the interpretation and generalisability of findings across the evidence base. Heterogeneity in intervention protocols represents a significant challenge for synthesis and clinical implementation. CFT formats ranged from brief six-session interventions to extended 20-session models, differing in delivery mode, session length, and therapeutic content. Most were delivered in group formats (typically six to eight participants) over 12 weekly sessions. Group settings may offer therapeutic benefits such as social support and opportunities to practise interpersonal compassion (Lawrence & Lee, 2014; Ashfield et al., 2021), but comparative studies of group versus individual delivery are lacking. This limits our understanding of optimal implementation strategies and clinical scalability.

Risk of bias was consistently high across the included studies. Common methodological concerns included a lack of participant masking, high attrition, selection bias in recruitment procedures, and potential co-intervention in some studies. These issues compromise the internal validity of the findings and reduce confidence in the reported effects. Many studies have also used single-arm designs without control groups, limiting the ability to attribute outcomes specifically to CFT.

Another critical limitation is the lack of follow-up data. Only six studies included follow-up data beyond immediate post-intervention timepoints, making it difficult to evaluate the long-term durability of CFT effects. This represents a significant gap, given that sustained therapeutic benefit is a key consideration for clinical implementation and cost-effectiveness analyses.

Demographic homogeneity remains a significant limitation affecting the generalisability of the evidence base. The pronounced gender imbalance and limited inclusion of non-Western populations and underrepresented clinical groups may obscure culturally specific experiences of compassion, shame, and self-criticism. This lack of diversity restricts the applicability of findings to broader clinical populations and highlights the need for future research to prioritise more representative sampling to better capture variations across gender, culture, and clinical presentation.

4.3. Appraisal of Review Processes

The review followed a rigorous and systematic methodology to enhance reliability and validity. It focused uniquely on outcomes integral to CFT theory. Grey literature was included to reduce publication bias; however, only one unpublished dissertation was identified (McLean et al., 2022), which was later published. The published, peer-reviewed version was included to minimise bias. The search was limited to English-language studies due to practical constraints. The English-language limitation may have introduced bias, affecting the representativeness of the included evidence (Dwan et al., 2013).

While the review applied rigorous inclusion criteria to ensure relevance to clinical practice, these criteria may have inadvertently limited the representation of diverse populations. For example, a study examining CFT for transgender and gender non-conforming individuals was excluded due to the absence of formal clinical classification measures, despite addressing substantial psychological distress (Sessions et al., 2023). Studies involving individuals with substance use or misuse as a primary diagnosis were also excluded. Although such populations could reasonably be included in future reviews, substance use fell outside the defined scope of this review. These exclusions limit the generalisability of the findings and prohibit important perspectives. This reflects a broader methodological tension in systematic reviews: balancing clear inclusion standards with the need for comprehensive and inclusive representation (Page et al., 2021).

To minimise selection bias, an anonymous, independent reviewer assessed a randomly selected subset of records (10% for screening; 20% for quality assessment), achieving perfect agreement in both cases ($\kappa = 1.00$). The quality assessment used the validated EPHPP tool (Thomas et al., 2004), which has demonstrated content validity and test–retest reliability (Armijo-Olivo et al., 2010; Deeks et al., 2003). However, as the primary investigator conducted most of the review, there remains potential for confirmation bias. This review followed PRISMA guidelines, was pre-registered with PROSPERO, and adhered to SWiM guidance for methodological clarity and transparency.

4.4. Implications

4.4.1. Practice

CFT interventions demonstrate the strongest potential for improving self-compassion and reducing self-criticism in clinical populations. Significant improvements were observed in relatively short group sessions, highlighting the potential for efficient clinical implementation. However, further research is needed to understand the efficacy of CFT in individual formats. Clinicians seeking to address self-criticism, shame, and compassion-related difficulties may find CFT techniques valuable. Additionally, therapists may wish to use the outcome measures included in the reviewed studies to evaluate intervention efficacy.

4.4.2. Research

Future research should address several gaps. First, recruitment should prioritise underrepresented populations to enhance generalisability and ensure equitable access to compassion-based interventions. This includes participants across all ages, gender identities and sexual orientations, diverse racial/ethnic backgrounds and socioeconomic

levels, people of all abilities and neurotypes, those with substance use disorders, and non-Western samples. Second, standardised measures should be used across studies to facilitate comparison and synthesis. For example, the CEAS (Gilbert et al., 2017) for compassion flows and the EISS (Ferreira et al., 2022) for shame states. Third, more rigorous research designs are needed to assess long-term impacts and compare group versus individual delivery formats. Fourth, identifying the specific CFT techniques that drive change could refine therapeutic delivery. Finally, standardised measures of treatment fidelity are needed to assess therapist adherence and participant engagement with compassion practices between sessions. Currently, the field lacks validated tools to evaluate whether CFT is being delivered as intended or to measure the extent to which participants incorporate compassion exercises into their daily lives. Understanding these factors would help clarify which components of CFT contribute most significantly to outcomes and whether the degree of participant engagement with compassion practices mediates therapeutic benefits.

4.4.3. Policy

CFT cannot yet be confidently included in best practice guidelines without higher-quality evidence addressing current limitations. Continued funding for methodologically robust CFT research is crucial, as is increased awareness of CFT training opportunities. Policymakers should take a cautiously optimistic approach while supporting ongoing scientific investigation into this promising therapeutic approach.

5. Conclusions

This systematic review found that CFT effectively improves self-compassion and reduces self-criticism in clinical populations, with most studies reporting medium to large effect sizes. It provides promising evidence that CFT may reduce external shame, with consistent findings across studies measuring this outcome. However, internal shame remains understudied despite its theoretical importance. The review highlights significant measurement gaps in the interpersonal compassion flows (compassion to others and receiving compassion from others) with mixed findings from the few studies examining these outcomes. These patterns suggest that while CFT shows robust effects on self-directed processes, its impact on relational aspects of compassion requires further investigation. Future research should address methodological limitations, compare delivery formats, and use consistent, theory-driven outcome measures. Clinicians should consider CFT for targeting shame and self-criticism, while policymakers should recognise CFT's emerging potential within clinical populations.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/bs15081031/s1>, Search Strategy.

Author Contributions: Conceptualisation, N.B. and K.A.; methodology, N.B.; validation, K.A.; formal analysis, N.B.; investigation, N.B.; data curation, N.B.; writing—original draft preparation, N.B.; writing—review and editing, K.A.; visualisation, N.B.; supervision, K.A.; project administration, N.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This systematic review followed PRISMA guidelines (Page et al., 2021) and was pre-registered with PROSPERO (CRD42023441154).

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon request. The search strategy, inclusion/exclusion criteria, and quality assessment tool are available in Supplementary Materials.

Acknowledgments: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The review was primarily conducted and submitted as a doctoral thesis at Royal Holloway University of London in 2024 and was then further refined post-qualification. The authors wish to thank Sophie Wilson (Clinical Psychologist) for independent review contributions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Altavilla, A., & Strudwick, A. (2022). Age inclusive compassion-focused therapy: A Pilot group evaluation. *International Journal of Cognitive Therapy*, 15, 209–230. [CrossRef]
- Armijo-Olivo, S., Stiles, C. R., Hagen, N. A., Biondo, P. D., & Cummings, G. G. (2010). Assessment of study quality for systematic reviews: A comparison of the cochrane collaboration risk of bias tool and the effective public health practice project quality assessment tool: Methodological research. *Journal of Evaluation in Clinical Practice*, 18(1), 12–18. [CrossRef] [PubMed]
- Arntz, A., van den Hoorn, M., Cornelis, J., Verheul, R., van den Bosch, W. M. C., & de Bie, A. J. H. T. (2003). Reliability and validity of the borderline personality disorder severity index. *Journal of Personality Disorders*, 17(1), 45–59. [CrossRef]
- Asano, K., Tsuchiya, M., Okamoto, Y., Ohtani, T., Sensui, T., Masuyama, A., Isato, A., Shoji, M., Shiraishi, T., Shimizu, E., Irons, C., & Gilbert, P. (2022). Benefits of group compassion-focused therapy for treatment-resistant depression: A pilot randomized controlled trial. *Frontiers in Psychology*, 13(13), 903842. [CrossRef]
- Ashfield, E., Chan, C., & Lee, D. (2021). Building “a compassionate armour”: The journey to develop strength and self-compassion in a group treatment for complex post-traumatic stress disorder. *Psychology and Psychotherapy: Theory, Research and Practice*, 94, 286–303. [CrossRef]
- Beck, A. T., Steer, R. A., Beck, J. S., & Newman, C. F. (1993). Hopelessness, depression, suicidal ideation, and clinical diagnosis of depression. *Suicide & Life-Threatening Behavior*, 23(2), 139–145.
- Beckham, G. (2016, February 3). *Cohen's d and hedges' g excel calculator*. GeorgeBeckham.com (original site no longer available). Available online: <https://medinterholguin2023.sld.cu/index.php/medintern/2023/paper/downloadSuppFile/33/3> (accessed on 15 November 2024).
- Behrouzian, F., Boostani, H., Mousavi Asl, E., Sadrizadeh, N., & Nemati Jole Karan, S. (2024). The effectiveness of compassion-focused therapy on patients with bipolar disorder, a randomized clinical trial. *Galen Medical Journal*, 13, e3359. [CrossRef]
- Boersma, K., Håkanson, A., Salomonsson, E., & Johansson, I. (2015). Compassion focused therapy to counteract shame, self-criticism and isolation. A replicated single case experimental study for individuals with social anxiety. *Journal of Contemporary Psychotherapy*, 45(2), 89–98. [CrossRef]
- Braehler, C., Harper, J., & Gilbert, P. (2013). Compassion focused group therapy for recovery after psychosis. In C. Steel (Ed.), *Cognitive behavior therapy for schizophrenia: Evidence-based interventions and future directions* (pp. 236–266). Blackwell.
- Callow, T. J., Moffitt, R. L., & Neumann, D. L. (2021). External shame and its association with depression and anxiety: The moderating role of self-compassion. *Australian Psychologist*, 56(1), 70–80. [CrossRef]
- Clapton, N. E., Williams, J., Griffith, G. M., & Jones, R. S. (2018). “Finding the person you really are . . . on the inside”: Compassion focused therapy for adults with intellectual disabilities. *Journal of Intellectual Disabilities*, 22(2), 135–153. [CrossRef]
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge.
- Cook, D. R. (1996). Empirical studies of shame and guilt: The Internalized Shame Scale. In D. L. Nathanson (Ed.), *Knowing feeling: Affect, script, and psychotherapy* (pp. 132–165). W. W. Norton & Company.
- 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), 1–16. [CrossRef]
- Deeks, J. J., Dinnes, J., D’Amico, R., Sowden, A. J., Sakarovich, C., Song, F., Petticrew, M., & Altman, D. G. (2003). Evaluating non-randomised intervention studies. *Health Technology Assessment*, 7(27), 1–173. [CrossRef]
- Dwan, K., Gamble, C., Williamson, P. R., & Kirkham, J. J. (2013). Systematic review of the empirical evidence of study publication bias and outcome reporting bias—An updated review. *PLoS ONE*, 8(7), e66844. [CrossRef] [PubMed]
- Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., & McGrath, G. (2000). CORE: Clinical outcomes in routine evaluation. *Journal of Mental Health*, 9(3), 247–255. [CrossRef]
- Ferreira, C., Moura-Ramos, M., Matos, M., & Galhardo, A. (2022). A new measure to assess external and internal shame: Development, factor structure and psychometric properties of the external and internal shame scale. *Current Psychology*, 41(4), 1892–1901. [CrossRef]
- First, M. B., & Gibbon, M. (2004). The structured clinical interview for DSM-IV axis I disorders (SCID-I) and the structured clinical interview for DSM-IV axis II disorders (SCID-II). In M. J. Hilsenroth, & D. L. Segal (Eds.), *Comprehensive handbook of psychological assessment* (Vol. 2. Personality assessment, pp. 134–143). John Wiley & Sons, Inc.

- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1996). *Structured clinical interview for DSM-IV axis I disorders, clinician version (SCID-CV)*. American Psychiatric Press.
- Fox, J., Cattani, K., & Burlingame, G. M. (2021). Compassion focused therapy in a university counseling and psychological services center: A feasibility trial of a new standardized group manual. *Psychotherapy Research*, 31(4), 419–431. [\[CrossRef\]](#) [\[PubMed\]](#)
- Gharraee, B., Zahedi Tajrishi, K., Ramezani Farani, A., Bolhari, J., & Farahani, H. (2018). A randomized controlled trial of compassion focused therapy for social anxiety disorder. *Iranian Journal of Psychiatry and Behavioral Sciences*, 12(4), 80945. [\[CrossRef\]](#)
- Gilbert, P. (2000). The relationship of shame, social anxiety and depression: The role of the evaluation of social rank. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice*, 7(3), 174–189.
- Gilbert, P. (2001). Evolution and social anxiety. *Psychiatric Clinics of North America*, 24(4), 723–751. [\[CrossRef\]](#)
- Gilbert, P. (2005). *Compassion: Conceptualisations, research and use in psychotherapy*. Routledge.
- Gilbert, P. (2009). Introducing compassion-focused therapy. *Advances in Psychiatric Treatment*, 15(3), 199–208. [\[CrossRef\]](#)
- Gilbert, P. (2010). *Compassion focused therapy: Distinctive features*. Routledge. [\[CrossRef\]](#)
- Gilbert, P. (2014). The origins and nature of compassion focused therapy. *British Journal of Clinical Psychology*, 53(1), 6–41. [\[CrossRef\]](#)
- Gilbert, P. (2015). The evolution and social dynamics of compassion. *Social and Personality Psychology Compass*, 9(6), 239–254. [\[CrossRef\]](#)
- Gilbert, P. (2017). *Compassion focused group therapy for college counseling centers*. Compassionate Mind Foundation.
- Gilbert, P. (2020). Compassion: From its evolution to a psychotherapy. *Frontiers in Psychology*, 11, 586161. [\[CrossRef\]](#) [\[PubMed\]](#)
- Gilbert, P., Basran, J. K., Raven, J., Gilbert, H., Petrocchi, N., Cheli, S., Rayner, A., Hayes, A., Lucre, K., Minou, P., Giles, D., Byrne, F., Newton, E., & McEwan, K. (2022). Compassion focused group therapy for people with a diagnosis of bipolar affective disorder: A feasibility study. *Frontiers in Psychology*, 13, 841932. [\[CrossRef\]](#) [\[PubMed\]](#)
- Gilbert, P., Catarino, F., Duarte, C., Matos, M., Kolts, R., Stubbs, J., Ceresatto, L., Duarte, J., Pinto-Gouveia, J., & Basran, J. (2017). The development of compassionate engagement and action scales for self and others. *Journal of Compassionate Health Care*, 4(1), 4. [\[CrossRef\]](#)
- Gilbert, P., & Choden. (2013). *Mindful compassion: Using the power of mindfulness and compassion to transform our lives*. Robinson.
- Gilbert, P., Clark, M., Hempel, S., Miles, J. N. V., & Irons, C. (2004). Criticising and reassuring oneself: An exploration of forms, styles and reasons in female students. *British Journal of Clinical Psychology*, 43, 31–50. [\[CrossRef\]](#)
- Gilbert, P., & Irons, C. (2004). A pilot exploration of the use of compassionate images in a group of self-critical people. *Memory*, 12(4), 507–516. [\[CrossRef\]](#) [\[PubMed\]](#)
- Gilbert, P., Kirby, J., & Petrocchi, N. in press. *Compassion focused therapy: Group therapy guidance*. Routledge.
- 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*, 13(6), 353–379. [\[CrossRef\]](#)
- Gilbert, P., & Simos, G. (Eds.). (2022). *Compassion focused therapy: Clinical practice and applications*. Routledge.
- Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., Heninger, G. R., & Charney, D. S. (1989). The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. *Archives of General Psychiatry*, 46(11), 1006–1011. [\[CrossRef\]](#)
- Goss, K., Gilbert, P., & Allan, S. (1994). An exploration of shame measures—I: The other as shamer scale. *Personality and Individual Differences*, 17(5), 713–717. [\[CrossRef\]](#)
- Grodin, J., Clark, J. L., Kolts, R., & Lovejoy, T. I. (2019). Compassion focused therapy for anger: A pilot study of a group intervention for veterans with PTSD. *Journal of Contextual Behavioral Science*, 13, 27–33. [\[CrossRef\]](#)
- Hedges, L. V. (1981). Distribution theory for glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6(2), 107–128. [\[CrossRef\]](#)
- Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2012). A standardized mean difference effect size for single case designs. *Research Synthesis Methods*, 3(3), 224–239. [\[CrossRef\]](#)
- Heriot-Maitland, C., Vidal, J. B., Ball, S., & Irons, C. (2014). A compassionate-focused therapy group approach for acute inpatients: Feasibility, initial pilot outcome data, and recommendations. *British Journal of Clinical Psychology*, 53(1), 78–94. [\[CrossRef\]](#)
- Johannsen, M., Schlander, C., Farver-Vestergaard, I., Lundorff, M., Wellnitz, K. B., Komischke-Konnerup, K. B., & O'Connor, M. (2022). Group-based compassion-focused therapy for prolonged grief symptoms in adults—Results from a randomized controlled trial. *Psychotherapy Research*, 314, 114683. [\[CrossRef\]](#)
- Judge, L., Clegghorn, A., McEwan, K., & Gilbert, P. (2012). An exploration of group-based compassion focused therapy for a heterogeneous range of clients presenting to a community mental health team. *International Journal of Cognitive Therapy*, 5(4), 420–429. [\[CrossRef\]](#)
- Kirby, J. N., & Gilbert, P. (2019). Commentary regarding Wilson et al. (2018) “Effectiveness of ‘self-compassion’ related therapies: A systematic review and meta-analysis.” All is not as it seems. *Mindfulness*, 10(6), 1006–1016. [\[CrossRef\]](#)
- 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. [\[CrossRef\]](#) [\[PubMed\]](#)

- Kolts, R. L. (2015). *True strength: A compassion focused therapy approach for working with anger*. Available online: <https://contextualconsulting.co.uk/wp-content/uploads/2023/07/True-Strength-Manual-Anger-2016.pdf> (accessed on 15 November 2024).
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606–613. [CrossRef]
- Laithwaite, H., O'Hanlon, M., Collins, P., Doyle, P., Abraham, L., Porter, S., & Gumley, A. (2009). Recovery after psychosis (RAP): A compassion focused programme for individuals residing in high security settings. *Behavioural and Cognitive Psychotherapy*, 37(5), 511–526. [CrossRef] [PubMed]
- Lambert, M. J., Gregersen, A. T., & Burlingame, G. M. (2004). The outcome questionnaire-45. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcomes assessment: Instruments for adults* (3rd ed., pp. 191–234). Lawrence Erlbaum Associates Publishers.
- Lawrence, V. A., & Lee, D. (2014). An exploration of people's experiences of compassion-focused therapy for trauma, using interpretative phenomenological analysis. *Clinical Psychology & Psychotherapy*, 21(6), 495–507. [CrossRef]
- Leaviss, J., & Uttley, L. (2015). Psychotherapeutic Benefits of compassion-focused therapy: An Early Systematic Review. *Psychological Medicine*, 45(5), 927–945. [CrossRef]
- Lee, D., & James, S. (2012). *The Compassionate mind approach to recovering from trauma* (1st ed.). Robinson.
- Lee, D., Scragg, P., & Turner, S. (2001). The role of shame and guilt in traumatic events: A clinical model of shame-based and guilt-based PTSD. *British Journal of Medical Psychology*, 74(4), 451–466. [CrossRef] [PubMed]
- Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the depression anxiety stress scales (DASS) with the beck depression and anxiety inventories. *Behaviour Research and Therapy*, 33(3), 335–343. [CrossRef] [PubMed]
- 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. [CrossRef] [PubMed]
- McLean, L., Steindl, S. R., & Bambling, M. (2022). Compassion focused group therapy for adult female survivors of childhood sexual abuse: A preliminary investigation. *Mindfulness*, 13, 1144–1157. [CrossRef]
- McManus, J., Tsivos, Z., Woodward, S., Fraser, J., & Hartwell, R. (2018). Compassion focused therapy groups: Evidence from routine clinical practice. *Behaviour Change*, 35(3), 167–173. [CrossRef]
- Methley, A. M., Campbell, S., Chew-Graham, C., McNally, R., & Cheraghi-Sohi, S. (2014). PICO, PICOS and SPIDER: A comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Services Research*, 14(1), 1–10. [CrossRef]
- Millard, L. A., Wan, M. W., Smith, D. M., & Wittkowski, A. (2023). The effectiveness of compassion focused therapy with clinical populations: A systematic review and meta-analysis. *Journal of Affective Disorders*, 326, 168–192. [CrossRef]
- Naismith, I., Ripoll, K., & Pardo, V. M. (2021). Group compassion-based therapy for female survivors of intimate-partner violence and gender-based violence: A pilot study. *Journal of Family Violence*, 36, 175–182. [CrossRef]
- Neff, K. (2003). The development and validation of a scale to measure self-compassion. *Self and Identity*, 2(3), 223–250. [CrossRef]
- Noorbala, F., Borjali, A., Ahmadian-Attari, M. M., & Noorbala, A. A. (2013). Effectiveness of compassionate mind training on depression, anxiety, and self-criticism in a group of Iranian depressed patients. *Iranian Journal of Psychiatry*, 8(3), 113–117.
- Otsubo, T., Tanaka, K., Koda, R., Shinoda, J., Sano, N., Tanaka, S., Aoyama, H., Mimura, M., & Kamijima, K. (2005). Reliability and validity of Japanese version of the mini-international neuropsychiatric interview. *Psychiatry and Clinical Neurosciences*, 59(5), 517–526. [CrossRef]
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., & McGuinness, L. A. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *British Medical Journal*, 372, n71. [CrossRef]
- Petrocchi, N., Cosentino, T., Pellegrini, V., Femia, G., D'Innocenzo, A., & Mancini, F. (2021). Compassion-focused group therapy for treatment-resistant OCD: Initial evaluation using a multiple baseline design. *Frontiers in Psychology*, 11, 594277. [CrossRef]
- Pol, S. M., de Jong, A., Trompetter, H., Bohlmeijer, E. T., & Chakhssi, F. (2024). Effectiveness of compassion-focused therapy for self-criticism in patients with personality disorders: A multiple baseline case series study. *Personality and Mental Health*, 18(1), 69–79. [CrossRef] [PubMed]
- Pol, S. M., Pots, W. T. M., & Bohlmeijer, E. (2020). *Compassion-focused therapy: Group protocol*. University of Twente.
- Prigerson, H. G., Horowitz, M. J., Jacobs, S. C., Parkes, C. M., Aslan, M., Goodkin, K., Raphael, B., Marwit, S. J., Wortman, C., Neimeyer, R. A., Bonanno, G., Block, S. D., Kissane, D., Boelen, P., Maercker, A., Litz, B. T., Johnson, J. G., First, M. B., & Maciejewski, P. K. (2009). Prolonged grief disorder: Psychometric validation of criteria proposed for DSM-V and ICD-11. *PLoS Medicine*, 6(8), e1000121. [CrossRef]
- 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. [CrossRef]

- Rogier, G., Muzi, S., Morganti, W., & Pace, C. S. (2023). Self-criticism and attachment: A systematic review and meta-analysis. *Personality and Individual Differences*, 214, 112359. [\[CrossRef\]](#)
- Rosenthal, R. (1994). Science and ethics in conducting, analyzing, and reporting psychological research. *Psychological Science*, 5(3), 127–134. [\[CrossRef\]](#) [\[PubMed\]](#)
- Santor, D. A., Zuroff, D. C., Mongrain, M., & Fielding, A. (1997). Validating the mcgill revision of the depressive experiences questionnaire. *Journal of Personality Assessment*, 69(1), 164–182. [\[CrossRef\]](#)
- Savari, Y., Mohagheghi, H., & Petrocchi, N. (2021). A preliminary investigation on the effectiveness of compassionate mind training for students with major depressive disorder: A randomized controlled trial. *Mindfulness*, 12, 1159–1172. [\[CrossRef\]](#)
- Sessions, L., Pipkin, A., Smith, A., & Shearn, C. (2023). Compassion and gender diversity: Evaluation of an online compassion-focused therapy group in a gender service. *Psychology & Sexuality*, 14(3), 528–541. [\[CrossRef\]](#)
- Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder. *Archives of Internal Medicine*, 166(10), 1092–1097. [\[CrossRef\]](#) [\[PubMed\]](#)
- Stroud, J., & Griffiths, C. (2021). An evaluation of compassion-focused therapy within adult mental health inpatient settings. *Psychology and Psychotherapy: Theory, Research and Practice*, 94(3), 587–602. [\[CrossRef\]](#)
- Thomas, B. H., Ciliska, D., Dobbins, M., & Micucci, S. (2004). A process for systematically reviewing the literature: Providing the research evidence for public health nursing interventions. *Worldviews on Evidence-Based Nursing*, 1(3), 176–184. [\[CrossRef\]](#) [\[PubMed\]](#)
- Thompson, R., & Zuroff, D. C. (2004). The levels of self-criticism scale: Comparative self-criticism and internalized self-criticism. *Personality and Individual Differences*, 36(2), 419–430. [\[CrossRef\]](#)
- Vanderheiden, E., & Mayer, C. H. (Eds.). (2024). Gender-specific facets of shame: Exploring a resource within and between cultures. In *Shame and gender in transcultural contexts*. Springer. [\[CrossRef\]](#)
- Wakelin, K. E., Perman, G., & Simonds, L. M. (2021). Effectiveness of self-compassion related interventions for reducing self-criticism: A systematic review and meta-analysis. *Clinical Psychology & Psychotherapy*, 29(1), 1–25. [\[CrossRef\]](#)
- Weathers, F. W., Marx, B. P., Friedman, M. J., & Schnurr, P. P. (2014). Posttraumatic stress disorder in DSM-5: New criteria, new measures, and implications for assessment. *Psychological Injury and Law*, 7(2), 93–107. [\[CrossRef\]](#)
- Weiss, D. S., & Marmar, C. R. (1997). The impact of event scale—Revised. In J. P. Wilson, & T. M. Keane (Eds.), *Assessing psychological trauma and PTSD* (pp. 399–411). The Guilford Press.
- Welford, M. (2012). *The compassionate mind approach to building your self-confidence using compassion focused therapy*. Robinson.
- Wilson, A. C., Mackintosh, K., Power, K., & Chan, S. W. Y. (2018). Effectiveness of self-compassion related therapies: A systematic review and meta-analysis. *Mindfulness*, 10(6), 979–995. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.