

RESEARCH ARTICLE

The association between self-compassion and treatment outcomes: Session-level and treatment-level effects

Lior Galili-Weinstock¹  | Roei Chen¹ | Dana Atzil-Slonim¹ | Eran Bar-Kalifa² | Tuvia Peri¹ | Eshkol Rafaeli¹

¹Bar-Ilan University

²Ben Gurion University of the Negev

Correspondence

Lior Galili-Weinstock, Department of Psychology, Bar-Ilan University, Ramat Gan, 52900, Israel.

Email: galili.lior@gmail.com

Abstract

Objective Self-compassion (SC) has been consistently linked to less psychopathology; however, the link between changes in client's SC levels and psychotherapy outcomes has yet to be explored.

Method Clients at a university-based community clinic completed SC and outcome measures session by session (N = 112) as well as pre- to posttreatment (N = 70).

Results Increases in clients' SC levels across the entire therapeutic process were associated with improvement in all posttreatment outcomes. Additionally, session-to-session increases in SC levels predicted improved symptoms and functioning at the session level; these effects were significant above and beyond the effects of the therapeutic alliance.

Conclusion The results of the current study highlight SC as a possible process variable in psychotherapy.

KEYWORDS

outcome, process, psychotherapy, self-compassion

1 | INTRODUCTION

Self-compassion (SC) is defined as healthy self-to-self relating, which involves a nonjudgmental attitude toward one's pain, inadequacies, and failures, with the perception that these are part of the greater human experience (Neff, 2003a). Growing evidence suggests that SC is associated with numerous psychological strengths as well as with psychological, cognitive, and emotional well-being (for a recent meta-analysis, see Zessin, Dickhäuser, & Garbade, 2015). Greater SC has consistently been linked to less psychopathology (Barnard & Curry, 2011) and specifically with lower levels of depression, anxiety, and stress (for a meta-analytic review, see MacBeth & Gumley, 2012).

Several authors have suggested that SC is a central mechanism of change in psychotherapy that leads to improved therapeutic outcomes (e.g., Germer & Neff, 2013; Gilbert, 2010). However, the link between changes in clients' SC

levels and treatment outcomes has been investigated primarily in the circumscribed context of SC-enhancement protocols (Barnard & Curry, 2011). Most relevant studies have been pilot studies examining the benefits of SC enhancement in small samples and with a limited number of assessment time points (typically two to three). Thus, the available data (e.g., Germer & Neff, 2013; Schanche, Stiles, McCullough, Svartberg, & Nielsen, 2011) refer to changes from pre- to postintervention but tell us little about changes and fluctuations from session to session. The current study aims to explore SC in the context of naturalistic psychotherapy processes and is the first to trace changes in SC and outcome measures using session-by-session monitoring. This methodology allows us to examine the association between changes in SC and changes in symptoms and functioning on both session (session-to-session symptoms and functioning) and treatment (pre- to posttreatment changes) levels.

1.1 | What is self-compassion?

SC is a core concept in Buddhist thinking (e.g., Brach, 2004; Salzberg, 1997) that has recently garnered much research attention in social, personality, and clinical psychology (Barnard & Curry, 2011). SC has been approached from several theoretical perspectives highlighting its role in affect regulation (Gilbert, 2010), attention and intentionality (Dalai Lama, 2001), and healthy self-to-self relating (Neff, 2003a). These models reflect subtle but important differences; however, all of them predict that SC would be associated with improved wellbeing and reduced emotional distress (MacBeth & Gumley, 2012).

Neff's (2003a) operationalization of SC comprises three main elements: (a) self-kindness, (b) common humanity, and (c) mindfulness. Although these elements may be experienced separately, they typically interact, as each causes the others to develop and grow. *Self-kindness* means treating oneself gently in the midst of suffering and is the opposite of *self-judgment*. *Common humanity* indicates the ability to recognize that suffering is not a private experience, as all humans fail, make mistakes, or engage in dysfunctional behaviors. This view stands in opposition to *isolation*, in which one loses sight of the larger human picture and focuses primarily on his/her own weakness or problems. Finally, *mindfulness* is a state of mind that allows individuals to observe and describe their thoughts and feelings without becoming overly engaged in them. Mindfulness stands in contrast to *overidentification* and rumination, and it represents the ability to experience things as they occur in the present moment without holding onto them or actively avoiding them (Hayes, Strosahl, & Wilson, 2002).

1.2 | Self-compassion in different psychotherapy models

Despite originating in Eastern philosophy and culture, the construct of SC, or some of its components, is consistent with the work of contemporary Western psychologists from a variety of theoretical orientations (e.g., Germer & Neff, 2013; Gilbert, 2010.; Greenberg & Watson, 2006; Young, Klosko, & Wieshar, 2003). In the psychoanalytic tradition, self-criticism or self-hatred is conceptualized in the context of an internalized parent-child relationship or childhood experiences of punishment (e.g., Aronfreed, 1964; Scharff & Tsigounis, 2003).

Blatt (1974, 1995) described self-criticism as a dimension of psychological vulnerability characterized by a sense of failure to fulfill one's standards and by feelings of inferiority and guilt (Blatt, Quinlan, Chevron, McDonald, & Zuroff, 1982). Shahar (2001) highlighted the role of shame as a mediating factor between self-criticism and psychological vulnerability and described how self-critical individuals set unrealistic goals in order to cope with intense feelings of shame and inadequacy. This dynamic creates a vicious cycle that exacerbates self-degradation and leads to various psychopathologies. These models accord with the concept of SC, specifically the elements of self-kindness and common humanity, which may serve as antidotes to self-criticism, isolation, and shame.

1.3 | Previous research on SC in nonclinical and clinical populations

Research on SC has primarily focused on its psychological and behavioral correlates in nonclinical populations. This large body of research has demonstrated that SC is positively linked with numerous psychological strengths as well as life satisfaction and feelings of social connectedness (e.g., Heffernan, Quinn Griffin, McNulty, & Fitzpatrick, 2010;

Hollis-Walker & Colosimo, 2011; Neff, Kirkpatrick, & Rude, 2007). Previous studies have suggested that SC may act as a buffer against psychological stressors (Gilbert, 2010) and promote better adjustment after personal challenges (Neff, Hsieh, & Dejitterat, 2005; Sbarra, Smith, & Mehl, 2012). SC has been hypothesized to promote adaptive emotion regulation (e.g., Neff et al., 2005), and several studies have reported negative links between SC and difficulties in emotional regulation (e.g., Finlay-Jones, Rees, & Kane, 2015) or specific maladaptive emotion regulation strategies, such as rumination, thought suppression, and avoidance (Krieger, Altenstein, Baettig, Doerig, & Holtforth, 2013; Neff et al., 2007).

Recently, a small number of studies have examined SC in clinical populations (e.g., Ferreira, Matos, Duarte, & Pinto-gouveia, 2014; Krieger et al., 2013; Kuyken et al., 2010; Lockard, Hayes, Neff, & Locke, 2014; Vettese, Dyer, Li, & Wekerle, 2011). SC levels have consistently been lower in these populations than in nonclinical populations (Costa, Marôco, Pinto-Gouveia, Ferreira, & Castilho, 2015; Roemer et al., 2009; Vettese et al., 2011). Moreover, SC has been found to be negatively associated with the level of (a) symptomatic distress among participants diagnosed with major depression (Krieger et al., 2013) or eating disorders (Ferreira et al., 2014) and (b) fears of negative and positive evaluation among participants diagnosed with social anxiety (but not with the anxiety symptoms themselves; Werner et al., 2012).

SC has also been tied to decreased difficulties with emotion regulation among adults with generalized anxiety disorder (Roemer et al., 2009) and adolescents and young adults with substance use-related problems (Vettese et al., 2011). Finally, in an experiment involving a sample of clinically depressed individuals who underwent a negative mood induction, SC was found to serve as an effective emotion regulation strategy (Diedrich, Grant, Hofmann, Hiller, & Berking, 2014).

1.4 | The association between changes in SC and treatment outcomes

While research on SC has grown at an exponential rate in recent years (Neff, 2016), studies exploring SC in the psychotherapy context have lagged far behind. To date, the link between changes in patients' SC levels and treatment outcomes has been investigated primarily in pilot studies of SC-enhancement protocols tested in small samples and focused on broad (pre- to postintervention) changes. For example, Gilbert and Procter (2006) conducted a pilot study of compassion-focused therapy in a group of day treatment patients who suffered from high levels of self-criticism and shame. After the group intervention, participants showed significant decreases in depression, self-attacks, shame, and feelings of inferiority. Neff and Germer (2013) compared the effect of mindful self-compassion training (MST) to the results for a waitlist control group in a nonclinical sample. Following MST, participants demonstrated significant increases in SC, mindfulness, compassion for others, and life satisfaction, as well as decreases in depression, anxiety, stress, and emotional avoidance. These outcomes were maintained at 6-month and 1-year follow-ups.

SC has been discussed as a mechanism of change in mindfulness-based cognitive therapy (MBCT; Teasdale et al., 2000). In a study comparing MBCT with maintenance antidepressants, pre- to postintervention increases in mindfulness and SC mediated the effect of MBCT on depressive symptoms at the 15-month follow-up (Kuyken et al., 2010). SC was also found to play an important role in both cognitive therapy and affect phobia therapy for patients suffering from cluster C personality disorders (Schanche et al., 2011). Increases in SC levels from early to late in therapy were found to significantly predict pre- to posttherapy decreases in psychiatric symptoms, interpersonal problems, and personality pathology in both treatment modalities.

Perfectionism and self-criticism, which may be conceptualized as the opposite of the self-kindness component of SC, have been negatively associated with SC (Neff, 2003a). Thus, psychotherapy studies in which these constructs have been examined are also relevant. Several of these studies have examined the role of self-critical perfectionism in psychotherapy outcomes (e.g., Blatt, Zuroff, Bondi, Sanislow, & Pilkonis, 1998; Hawley, Ho, Zuroff, & Blatt, 2006). For example, pretreatment perfectionism appears to impede therapeutic progress, resulting in a weaker reduction in depression levels (Blatt et al., 1998). Interestingly, self-criticism seems to disrupt the effectiveness of diverse treatment modalities, including pharmacotherapy (Blatt, Quinlan, Pilkonis, & Shea, 1995).

The extant literature seems to support the possible role of SC as a change process, but direct tests of SC within psychotherapy remain scarce. The current study thus seeks to explore this role at both the treatment level (i.e., from pre- to posttherapy) and the session level (i.e., from one week to the next). We hypothesized that (a) increases in clients' SC from pre- to posttreatment would be associated with lower levels of posttreatment symptomology after adjusting for pretreatment symptom levels (*the treatment level hypothesis*); and (b) increases in clients' SC between sessions would be associated with better session-level outcomes (i.e., reduced symptoms and improved functioning) after adjusting for the clients' symptoms in the previous session (*the session level hypothesis*). Importantly, we hypothesized that changes in clients' SC levels would predict outcomes above and beyond the effect of the therapeutic alliance, a significant factor that has been consistently linked to treatment outcome (for a meta-analytic review see Horvath, Del Re, Flückiger, & Symonds, 2011)

2 | METHOD

2.1 | Participants and treatment

2.1.1 | Clients

The participants were adults currently undergoing psychotherapy at a major university outpatient clinic. Of the 167 clients who applied to treatment and agreed to participate in the study, 112 (67%) began treatment and were included in the analyses. Clients' diagnoses were established based on the Mini International Neuropsychiatric Diagnostic Interview (MINI 5.0; Sheehan et al., 1998) for Axis I *Diagnostic and Statistical Manual of Mental Disorders Fourth Edition* (DSM-IV; American Psychiatric Association, 1994) diagnoses. Trained psychologists, who received weekly group supervision by a senior clinician (TP), administered the MINI 5.0 in the intake meeting. All intake sessions were audiotaped, and an independent clinician (LGW) again sampled and rated a random 25% of the interviews. The mean kappa values of the Axis I diagnoses was excellent ($k = 0.95$). Moderate interrater agreement was found for panic and generalized anxiety disorders, whereas excellent agreement was found for all other disorders.

Of our total sample, 23.2% of the clients had a single diagnosis, 22.3% had two diagnoses, and 12.5% had three or more diagnoses. The most common diagnoses were for comorbid anxiety and affective disorders¹ (27.7%), followed by affective disorders (10.7%), anxiety disorders (9.8%), other comorbid disorders (6.3%), obsessive-compulsive disorder (2.7%), and PTSD (.89%). A sizable group of clients (42.0%) reported experiencing relationship concerns, academic/occupational stress, or other problems but did not meet criteria for Axis I diagnosis.

Of the 112 clients who began the study, 22 (19.6%) dropped out of therapy, with dropout defined as clients' unilateral decision to end treatment without agreement of the therapist (see Callahan, Aubuchon-Endsley, Borja, & Swift, 2009; Keijsers, Kampman, & Hoogduin, 2001; Westmacott, Hunsley, Best, Rumstein-McKean, & Schindler, 2010). In addition, 20 clients (17.9%) completed therapy but did not complete the posttreatment outcome questionnaires. Thus, although our session-by-session analyses used data from all 112 clients, our posttreatment outcome analyses used data from 70 (62.5%) clients.

2.1.2 | Therapists

The participating clients were assigned to therapists in an ecologically valid manner based on real-world issues such as therapist availability and caseload. A total of 66 therapists treated the clients. The therapists were in clinical training on different stages during their 2–5-year clinical training program: 30 therapists treated one client each, 30 treated two clients each, and six treated three to five clients each. The therapists were blind to the study hypotheses. Each therapist received 1 hour of individual supervision and 4 hours of group supervision on a weekly basis. All therapy sessions were audiotaped for use in supervision with senior clinicians. The individual and group supervision focused heavily on the review of audiotaped case material.

Individual psychotherapy consisted of once- or twice-weekly sessions of primarily psychodynamic psychotherapy organized, aided, and informed (but not prescribed) by a short-term psychodynamic psychotherapy treatment

model (Blagys & Hilsenroth, 2000; Shedler, 2010). The key features of this model are as follows: (a) focus on affect and the experience and expression of emotions; (b) exploration of attempts to avoid distressing thoughts and feelings; (c) identification of recurring themes and patterns; (d) emphasis on past experiences; (e) focus on interpersonal experiences; (f) emphasis on the therapeutic relationship; and (g) exploration of wishes, dreams or fantasies (Shedler, 2010).

Individual Treatment length was open ended; however, given the constraints of the university-based outpatient community clinic, which operates on an academic schedule, treatment length was often limited to 9–12 months. The mean treatment length was 22.1 sessions (standard deviation [SD] = 8.5, range = 5–47). A total of 2,472 sessions were available for analysis.

2.2 | Measures

2.2.1 | Treatment-level measures

2.2.2 | Self-Compassion Scale (SCS; Neff, 2003b)

This 26-item scale assesses six different aspects of SC, three of which are positive: (a) self-kindness (e.g., “I try to be understanding and patient towards those aspects of my personality I don't like”); (b) common humanity (e.g., “I try to see my failings as part of the human condition”); and (c) mindfulness (e.g., “When something painful happens I try to take a balanced view of the situation”). The other three aspects are negative: (d) self-judgment (e.g., “I'm disapproving and judgmental about my own flaws and inadequacies”; reverse scored); (e) isolation (e.g., “When I think about my inadequacies it tends to make me feel more separate and cut off from the rest of the world”); and (f) over-identification (e.g., “When I'm feeling down I tend to obsess and fixate on everything that's wrong”). These negative aspects are reverse coded.

Responses for all items are given on a 5-point scale ranging from 1 (*almost never*) to 5 (*almost always*). The SCS total scores range from 26 to 130. No clinical cutoff scores for the SCS have been published, but prior studies have reported mean scores for community samples to be around 80 with lower mean scores (60–70) for clinical populations (e.g., Castilho, Pinto-Gouveia, & Duarte, 2015; Krieger et al., 2013; Magnus, Kowalski, & McHugh, 2010; Neff, Whittaker, & Karl, 2017). The SCS has an appropriate factor structure, with a single overarching factor of “self-compassion” accounting for intercorrelations between subscales (but see Hayes, Lockard, Janis, & Locke, 2016), and the scale has demonstrated predictive, convergent, and discriminant validity (Neff, 2003b). The internal consistency in our sample was high ($\alpha = .91$).

2.2.3 | Outcome Questionnaire-45 (OQ-45; Lambert et al., 1996)

The OQ-45 is a self-report measure designed to assess patient outcomes over the course of therapy. The 45 items assess three primary dimensions: (a) subjective discomfort (e.g., anxiety and depression—“I feel blue”); (b) interpersonal relationships (e.g., “I feel lonely”); and (c) social role performance (e.g., “I have too many disagreements at work/school”). All 45 items can be aggregated to create total scores that range from 0 to 180, with higher scores reflecting poorer psychological functioning. The OQ-45 has been shown to have good internal consistency ($\alpha = .93$) and 3-week test-retest reliability ($r = .84$; Lambert et al., 1996). The high internal consistency was also replicated in our sample ($\alpha = .93$).

2.2.4 | Beck Depression Inventory (BDI-II; Beck, Steer, & Brown, 1996)

The BDI-II assesses the severity of depressive symptoms on the basis of 21 self-reported items. Respondents rate each statement on a scale ranging from 0 to 3, with higher rates representing more depressed affect. The total score is computed as the sum of all items (0–63). The internal consistency of the total score has been reported as good ($\alpha = .90$; Storch, Roberti, & Roth, 2004). The measure obtained high internal consistency in our sample as well ($\alpha = .90$).

2.2.5 | Difficulties in Emotion Regulation Scale (DERS; Gratz & Roemer, 2004)

The DERS is a 36-item measure that assesses typical levels of emotion dysregulation across six separate domains: (a) nonacceptance of negative emotions, (b) inability to engage in goal-directed behaviors when experiencing negative emotions, (c) difficulty in controlling impulsive behaviors when experiencing negative emotions, (d) limited access to emotion regulation strategies perceived as effective, (e) lack of emotional awareness, and (f) lack of emotional clarity. Respondents rate each statement on a 5-point Likert scale ranging from 1 (*almost never*, 0–10%) to 5 (*almost always*, 91–100%).

The DERS total scores range from 36 to 180. While there are no standardized clinical cutoffs for this measure, prior research indicated mean scores for nonpatient adults to be around 60 (e.g., Fox, Axelrod, Paliwal, Sleeper, & Sinha, 2007), with clinical samples evincing higher mean scores, ranging from approximately 80 to 127 (e.g., Fox, Hong, & Sinha, 2008; Gratz & Gunderson, 2006; Neacsiu, Eberle, Kramer, Wiesmann, & Linehan, 2014). The DERS has been proven to be sensitive to changes over time (Gratz & Gunderson, 2006), and it showed high internal consistency in our sample ($\alpha = .95$).

2.2.6 | Session-level measures

2.2.7 | Session-level self-compassion index

To monitor changes in patients' SC levels from session to session, we chose three SCS (Neff, 2003b) items, each representing a different positive SC subscale: (a) self-kindness ("When I had a hard time, I gave myself the caring and tenderness I needed"); (b) common humanity ("I tried to see my failings as part of the human condition"); and (c) mindfulness ("When something upset me I tried to keep my emotions in balance"). Patients were asked to rate each statement on a 5-point scale ranging from 1 (*almost never*) to 5 (*almost always*) with regards to the previous week. The three items were chosen based on a previous study that demonstrated their strong correlations with the SCS total score and with the individual subscale from which they were drawn (Raes, Pommier, Neff, & Van Gucht, 2011).

These results were replicated in our pilot sample of clients applying for treatment in our clinic ($N = 142$), and the following correlations were found between each item and the subscale from which it was drawn and with the full SCS, respectively: the self-kindness item ($r = .80, .65$, both $p < .001$); the common humanity item ($r = .78, .72$, both $p < .001$); and the mindfulness item ($r = .78, .56$, both $p < .001$). The between- and within-person reliabilities for the scale were computed using procedures outlined by Shrout and Lane (2012; see also Cranford et al., 2006), and these values were 0.79 and 0.77, respectively.

2.2.8 | Outcome Rating Scale (ORS; Miller, Duncan, Brown, Sparks, & Claud, 2003)

The ORS is a four-item visual analog scale developed as a brief alternative to the OQ-45. Three of its items assess changes in areas of client functioning that are widely considered valid indicators of progress in treatment: individual (or symptomatic) functioning, interpersonal relationships, and social role performance (work adjustment and quality of life). An additional item assesses overall functioning. The ORS has shown strong reliability ($\alpha = .87-.96$), and moderate correlations have been found between the ORS and the OQ-45 scores ($r = .59$). This correlation is at the expected level, given that 45 items were reduced to four (Miller et al., 2003). The visual analog scale is anchored at one end by the word *low* and at the other end by the word *high*, which are converted into scores from 0 to 10 and then summed to a total score ranging from 0 to 40, with higher scores indicating better functioning. The reliability levels in the current study were high (within = .91, between = 1.0).

2.2.9 | The Hopkins Symptom Checklist-Short Form (HSCL-11; Lutz, Tholen, Schürch, & Berking, 2006)

As an 11-item self-report inventory assessing symptomatic distress, the HSCL-11 is a brief version of the SCL-90-R (Derogatis, 1977). The items are rated on a 4-point Likert scale ranging from 1 (*not at all*) to 4 (*extremely*). The mean

TABLE 1 Sample demographic characteristics (N = 112)

Variable		
Gender (Female) (%)		58.9
Age (Mean± SD)		40.8± 13.65
Education (%)	Primary	1.9
	Secondary	35.5
	Higher education	62.6
Marital status (%)	Single	42.7
	Divorced or widowed	12.9
	Permanent relationship	44.4
Employment (%)	Unemployed	19.2
	Part time employment	24.2
	Full time Employment	56.6

Note. SD = standard deviation.

of the 11 items represents the client's level of global symptomatic distress during the preceding week. The score is highly correlated with the SCL-90-R's global severity index ($r = 0.91$) and has high internal consistency ($\alpha = .92$; Lutz et al., 2006). In the current study, the between- and within-person reliabilities for the scale were good (within = .84, between = .99).

2.2.10 | Client Working Alliance Inventory (WAI-SR; Hatcher & Gillaspay, 2006)

The 12-item short form of the WAI (Horvath & Greenberg, 1989) is based on Bordin's (1979) tripartite conceptualization of the client-therapist relationship, which includes agreement between the client and therapist on goals, the degree of concordance on tasks, and the strength of the therapeutic bond. Clients were asked to use a 7-point Likert scale to rate how accurately each item describes their current therapy experience. The WAI-SR has good reliability, with alpha coefficients for overall internal reliability ranging from .85 to .95. The reliability estimates of the subscales have also demonstrated fairly high alpha coefficients: .82 to .88 on the Task subscale, .82 to .87 on the Goal subscale, and .85 on the Bond subscale. The between- and within-person reliabilities found in our sample were high (within = .91, between = 1.0).

2.3 | Procedure

The study was conducted in a university-based outpatient clinic between August 2014 and August 2015. The study procedures were part of the routine monitoring battery in the clinic. Clients were asked to provide written consent to participate in the voluntary study and were told that they could choose to terminate their participation in the study at any time without jeopardizing their treatment. The study was conducted in compliance with the university ethical review board.

The SCS, OQ-45, BDI, and DERS questionnaires were administered to clients as part of the intake procedure (i.e., at pretreatment) and again following treatment termination. The clients completed the session-level questionnaires electronically using computers located in the clinic rooms. The session-level SC index, ORS, and HSCL-11 were completed before each therapeutic session; the WAI was completed after each therapeutic session.

3 | RESULTS

Clients' demographic characteristics are presented in Table 1. Pretreatment assessments included the OQ-45, BDI, DESR, and SCS. Means, standard deviations, and intercorrelations of all study variables are depicted in Table 2. As can

TABLE 2 Means, standard deviations and intercorrelations of study variables (N = 112)

		Mean \pm SD	OQ-45	BDI	DERS	SCS	HSCL	ORS
Pre-treatment measures	OQ-45	70.56 \pm 23.93						
	BDI	17.28 \pm 10.96	.831**					
	DERS	89.94 \pm 27.05	.69**	.69**				
	SCS	74.84 \pm 18.39	-.71**	-.62**	-.681**			
Session-level measures	HSCL	1.82 \pm 0.44	0.61***	0.62**	0.44***	-0.33**		
	ORS	25.14 \pm 6.52	-0.52***	-0.48***	-0.31***	0.22*	-0.56***	
	SC-Index	2.91 \pm 0.75	-0.35**	-0.31***	-0.32**	0.41***	-0.30**	-0.59***

Note. SD = standard deviation; OQ-45 = Outcome Questionnaire-45; BDI = Beck Depression Inventory; DERS = Difficulties in Emotion Regulation Scale; SCS = Self-Compassion Scale; HSCL = The Hopkins Symptom Checklist-Short Form; ORS = Outcome Rating Scale.

*** $p < .001$. ** $p < 0.01$. * $p < 0.05$.

be seen, the sample mean scores on pretreatment questionnaires indicated mild to moderate symptoms of impairment in psychological, social, and occupational functioning.

3.1 | The treatment-level hypothesis

To test whether the increase in clients' SC from pre- to posttreatment predicted better client outcomes at posttreatment, we conducted a hierarchical linear regression. In the first step, we entered the clients' pretreatment outcome, which allowed us to adjust for the initial level of the outcome variable and to treat the predicted outcome variable as a change score. Additionally, as noted above, we were interested in determining whether changes in clients' SC levels from pre- to posttreatment would predict outcomes above and beyond the effect of the therapeutic alliance. To do so, we computed the empirical Bayes estimates (Raudenbush & Bryk, 2002), which represent the clients' WAI slopes across sessions. These estimates (i.e., WAI slopes) were then entered to the regression.

In the second step, we entered the change in clients' SC level from pre- to posttreatment (i.e., Delta SC). Delta SC (mean [M] = .48, SD = 13.04) was calculated by subtracting clients' pretreatment SCS score from their posttreatment SCS score; hence, a higher Delta SC score represents an increase in a client's SC after treatment. This model was tested three times, once for each outcome variable²: (a) OQ-45, (b) BDI, and (c) DERS.

The results of these analyses are presented in Table 3. Consistent with our predictions, the Delta SCS score predicted the posttreatment outcome variables above and beyond the initial levels of the outcome scores. This result was true when predicting (a) posttreatment OQ-45 scores, R^2 change = .09, F change(1,65) = 18.79, $p < .001$, (b) posttreatment BDI scores, R^2 change = .12, F change(1,60) = 21.89, $p < .001$, and (c) posttreatment DERS scores, R^2 change = .05, F change(1,64) = 10.65, $p < .01$.

Given some concern about the use of difference scores based on two time-points, we computed a second model, replacing the delta SCS scores with an index utilizing the full set of repeated assessments and reflecting the trajectory of SC throughout treatment. To do so, we computed empirical Bayes estimates (Raudenbush & Bryk, 2002), which represent the slopes of the clients' SC across sessions (i.e., SC slope). This replaced delta SC in the second step of a hierarchical regression model, which was otherwise identical to the one described above. We entered the pretreatment outcome score as well as the WAI slope. This model was tested three times, once for each outcome measure.

The results of these analyses are presented in Tables 4. Again, Consistent with our predictions, the SC slope predicted the posttreatment outcome variables above and beyond the initial levels of the outcome scores and the therapeutic alliance for (a) posttreatment OQ-45 scores, R^2 change = .02, F change(1,74) = 4.14, $p < .05$, (b) posttreatment BDI scores, R^2 change = .05, F change(1,66) = 6.64, $p < .05$, and (c) posttreatment DERS scores, R^2 change = .01, F change(1,70) = 6.72, $p < .05$.

TABLE 3 Hierarchical linear regression predicting posttreatment scores

		Model	B	SE(B)	β	P	R ²	R ² change
OQ-45	1	(Constant)	11.93	5.29		.028	.60***	
		OQ-45 _(t1)	.69	.07	.75	.000		
		WAI slope	.06	.27	.02	.836		
	2	(Constant)	11.45	4.70		.018	.69***	.09***
		OQ-45 _(t1)	.71	.06	.79	.000		
		WAI slope	-.07	.25	-.02	.782		
BDI	1	(Constant)	2.734	1.692		.111	.53***	
		BDI _(t1)	.664	.082	.743	.000		
		WAI slope	-.101	.152	-.061	.508		
	2	(Constant)	2.83	1.46		.058	.66***	.13***
		BDI _(t1)	.67	.07	.75	.000		
		WAI slope	-.14	.13	-.06	.288		
DERS	1	(Constant)	25.49	6.39		.000	.61***	
		DERS _(t1)	.68	.07	.78	.000		
		WAI slope	-.04	.29	-.01	.896		
	2	(Constant)	23.65	5.99		.000	.66***	.06**
		DERS _(t1)	.70	.07	.81	.000		
		WAI slope	-.15	.27	-.04	.587		
		Delta SCS	-.44	.13	-.24	.002		

Note. SE = standard error; OQ-45 = Outcome Questionnaire-45; WAI = Client Working Alliance Inventory; SCS = Self-Compassion Scale; DERS = Difficulties in Emotion Regulation Scale.

***p < .001. **p < 0.01. *p < 0.05.

3.2 | The session-level hypothesis

Next, we tested whether the increase in clients' SC between sessions predicted better client outcomes at a given session when we controlled for the symptom level in the previous session. To control for the effect of the session-level therapeutic alliance, we included the previous session's therapeutic alliance score as a covariate³.

Because the session-level data have a hierarchical structure (sessions nested within clients nested within therapists), we used SAS PROC MIXED to estimate a multilevel model (MLM) for our predictions. We opted for two-level MLM (sessions nested within clients), rather than three-level MLM, because the level-three unconditional models for our outcome measures were either not significant (for HSCL as the outcome⁴) or entirely equal to zero (for ORS as the outcome).

Clients' SC change scores (i.e., Delta SC) were calculated by subtracting their SC levels at session_{s-1} from their SC levels at session_s; thus, higher scores indicated greater improvement in clients' SC. In addition, clients' outcome level at session_{s-1} was entered into the analysis to examine whether the improvement in SC predicted outcome levels while adjusting for earlier symptom levels (from session_{s-1}); this approach allowed us to treat the outcome level as a change score. As noted earlier, we also entered the therapeutic alliance (from session_{s-1}) as a covariate. The therapeutic alliance was centered around each client's mean to isolate the within-subject effects (Raudenbush & Byrk, 2001).

Finally, to account for the inherent effect of treatment progress on symptomatic distress, we entered the log₁₀ of the session number as a covariate. We opted to enter the log₁₀ of the session number in light of previous work demonstrating that symptomatic distress often changes over the course of therapy as a negatively accelerating function of

TABLE 4 Hierarchical linear regression predicting posttreatment scores

		Model	B	SE(B)	β	P	R ²	R ² change
OQ-45	1	(Constant)	13.12	4.96		.010	.59***	
		OQ-45 _(t1)	.69	.07	.77	.000		
		WAI slope	.00	.25	.00	.988		
	2	(Constant)	14.82	4.93		.004	.61***	.02*
		OQ-45 _(t1)	.67	.07	.75	.000		
		WAI slope	-.06	.25	-.02	.816		
BDI	1	(Constant)	2.73	1.69		.075	.51***	
		BDI _(t1)	.66	.08	.74	.000		
		WAI slope	-.10	.15	-.06	.568		
	2	(Constant)	2.96	1.64		.034	.55***	.05*
		BDI _(t1)	.65	.08	.72	.000		
		WAI slope	-.08	.14	-.05	.432		
DERS	1	(Constant)	24.83	5.89		.000	.63***	
		DERS _(t1)	.69	.06	.79	.000		
		WAI slope	-.02	.27	-.01	.940		
	2	(Constant)	25.84	5.69		.000	.66***	.03*
		DERS _(t1)	.69	.06	.79	.000		
		WAI slope	-.09	.26	-.03	.727		
		SC slope	-7.76	2.99	-.18	.012		

Note. OQ-45 = Outcome Questionnaire-45; BDI = Beck Depression Inventory; DERS = Difficulties in Emotion Regulation Scale; WAI = Client Working Alliance Inventory.

*** $p < .001$. ** $p < 0.01$. * $p < 0.05$.

the number of sessions, with the most rapid response occurring early in therapy (for a similar approach, see Lutz, Leon, Martinovich, Lyons, & Stiles, 2007). Two models were computed, once for each outcome measure (HSCL and ORS):

Level 1: Client's $\text{Outcome}_{cs} = \beta_{0c} + \beta_{1c} * \text{Outcome}_{(s-1)c} + \beta_{2c} * \text{mean-centered alliance}_{(s-1)c} + \beta_{3c} * \log_{10} \text{session number}_{sc} + \beta_{4c} * \Delta \text{SC} + e_{cs}$

Level 2: $\beta_{0c} = \gamma_{00} + u_{0c}$; $\beta_{1c} = \gamma_{10} + u_{1c}$; $\beta_{2c} = \gamma_{20} + u_{2c}$; $\beta_{3c} = \gamma_{30} + u_{3c}$; $\beta_{4c} = \gamma_{40} + u_{4c}$

The level-one equation modeled the outcome for client c in session s as a function of (a) the client's intercept, (b) the client's outcome level at session _{$s-1$} , (c) the mean-centered therapeutic alliance at session _{$s-1$} , (d) log₁₀ of the session number, (e) the client Δ SC from session _{$s-1$} to session _{s} , and finally (f) a level-one residual term. These baseline models were then modified based on nonsignificant random effects, as specified below.

For HSCL as the outcome, we entered the intercept (i.e., β_{0c}) as well as the fixed effects of each client (e.g., β_{1c}) as random effects because this improved the model fit, $\chi^2(2) = 175$, $p < .001$. For ORS as the outcome, we did not enter the therapeutic alliance as a random effect, as it did not have any significant variance at level two; all other effects were modeled as both fixed and random, which improved the model fit, $\chi^2(2) = 73$, $p < .001$.

The MLM results are shown in tables 5 and 6. As predicted, clients' improvement in SC level between sessions predicted their outcome at a given session above and beyond their outcome level in the previous session when we adjusted for the earlier session's therapeutic alliance and for the log₁₀ of the session number. This result held true for both HSCL and ORS as outcome measures.

TABLE 5 Multilevel model predicting HSCL scores

Parameter estimates	Estimate (SE)	Effect size ^a
Fixed effects		
Intercept (γ_{00})	1.16(.05)***	
Lagged HSCL (γ_{10})	0.39(.02)***	0.49
WAI	0.00(.00)	0.00
Log10 of session number	-0.09(.03)**	0.03
Δ SC (γ_{20})	-0.09(.02)***	0.19
Random effects		
Level 1 (sessions)		
Residual	0.08(.00)***	
Level 2 (clients)		
Intercept	0.01(.01)	
Lagged HSCL (u1c)	.01(.00)***	
WAI	0.00(.00)	
Log10 of session number	0.03(.01)***	
Δ SC (u2c)	.02(.00)***	
Model summary		
-2 Log L	1023.3	
# Estimated parameters	10	

Note. SE = standard error; HSCL = The Hopkins Symptom Checklist-Short Form; OQ-45 = Outcome Questionnaire-45; WAI = Client Working Alliance Inventory.

Effect sizes were calculated as semipartial R^2 (Edwards, Muller, Wolfinger, Qaqish, & Schabenberger, 2008).

* $p < .05$. ** $p < .01$. *** $p < .001$.

To estimate the global explained variance of our model, we calculated the correlation between the predicted and observed outcome values, which resulted in 57% and 59% of the variance explained in the whole model for HSCL and ORS, respectively (Peugh, 2010; Singer & Willett, 2003).

4 | DISCUSSION

The current study aimed to examine the association between SC and treatment outcomes in naturalistic psychotherapy. We hypothesized that increases in clients' SC would be associated with improved treatment outcomes at both the treatment and session levels. These predictions were fully supported by our results. First, we found increases in clients' SC levels throughout the therapeutic process to be associated with lower levels of posttreatment depression, overall symptomology, and emotional difficulties. A similar pattern of results emerged when we examined the association between SC and both session-level outcomes. Specifically, our results showed that session-to-session increases in SC levels predicted general functioning, as well as decreased levels of symptomology in the following session, even when we controlled for the outcome level in the previous session.

Interestingly, these effects were significant above and beyond the well-documented effects of the therapeutic alliance, a variable that has been consistently linked to therapy outcomes. Additionally, by using the log10 of the session number, we demonstrated that the effect of changes in clients' SC levels on session-level outcomes held even when controlling for the inherent log-linear effect of treatment progress.

Taken together, the current findings may be interpreted to suggest that treatment-related increases in clients' SC were tied to positive therapeutic outcomes. These findings accord with previous studies documenting the consistent association between SC and lower levels of psychopathology as well as with higher levels of well-being

TABLE 6 Multilevel model predicting ORS scores

Parameter estimates	Estimate (SE)	Effect size ^a
Fixed effects		
Intercept (γ_{00})	13.81(.07)***	
Lagged ORS (γ_{10})	0.38(.02)***	0.46
WAI	0.04(.01)**	0.00
Log10 of session number	1.98(.48)***	0.08
Δ SC (γ_{20})	1.52(.23)***	0.32
Random effects		
Level 1 (sessions)		
Residual	17.18(.06)***	
Level 2 (clients)		
Intercept	10.44(2.87)***	
Lagged ORS (u1c)	.01(.00)**	
WAI		
Log10 of session number	5.23(1.94) [†]	
Δ SC (u2c)	2.39(.06)***	
Model summary		
-2 Log L	11693.1	
# Estimated parameters	10	

Note. SE = standard error; ORS = Outcome Rating Scale; OQ-45 = Outcome Questionnaire-45; WAI = Client Working Alliance Inventory.

Effect sizes were calculated as semipartial R^2 (Edwards, Muller, Wolfinger, Qaqish, & Schabenberger, 2008).

[†] $p < .05$. ** $p < .01$. *** $p < .001$.

(for a meta-analysis review, see MacBeth & Gumley, 2012; Zessin et al., 2015). Similarly, our finding that increases in clients' SC were tied to lower levels of posttreatment emotional regulation difficulties is in line with previous studies reporting an association between SC and more adaptive emotional regulation (e.g., Finlay-Jones et al., 2015).

Our results are also consistent with work documenting an effect of specific SC-enhancement group interventions on SC and on symptomatic relief (e.g., Gilbert & Procter, 2006; Neff & Germer, 2013). The current study adds to this literature by demonstrating the negative link between SC and symptoms in the context of naturalistic (i.e., not SC-focused) psychodynamic psychotherapy. In addition, whereas previous studies have relied on smaller samples and fewer measurement points, this study is the first to trace changes in SC and outcomes over the entire course of treatment using ongoing session-by-session monitoring. By doing so, this study implements methodological recommendations made in recent years by prominent psychotherapy researchers, who have advocated for examining the direct associations between process and outcome variables as they fluctuate session by session throughout treatment (e.g., Crits-Christoph, Connolly Gibbons, & Mukherjee, 2013).

Our results lend some support to the theoretical view of SC as a curative factor (Gilbert, 2000; Neff, 2003a) that leads to better therapeutic outcomes. Notably, they demonstrate the significance of SC in a treatment model that is not explicitly focused on enhancement of SC. Still, the short-term psychodynamic framework does have certain key features that are likely to result in SC change. For example, psychodynamic therapy encourages clients to explore and express their full emotional range as therapists actively focus on clients' attempts to avoid distressing thoughts and feelings (Blagys, & Hilsenroth, 2000; Shedler, 2010). This process may enhance mindfulness and promote clients' ability to tolerate negative or unwanted emotions, gradually accepting them as part of the human experience.

Another key element of the psychodynamic therapy is the exploration of past experiences (especially early ones with attachment figures), linking them to the clients' present difficulties and shortcoming. This process changes clients' personal narrative and promotes self-understanding (Shedler, 2010). It is possible that such narrative work is helpful in reducing clients' self-blame regarding their personal difficulties, enhancing self-kindness, while at the same time promoting change.

Finally, the psychodynamic perspective highlights the therapeutic relationship and the therapist's empathic stance as a vehicle for the process of change (Blagys & Hilsenroth, 2000). Potentially, the therapist's nonjudgmental and supportive attitudes and behaviors provide a safe environment in which the client is able to disclose personal information without being afraid of rejection or criticism (Hawley et al., 2006). This process may help the client internalize the therapist's empathic attitude and develop more benign and compassionate self-to-self relating. While the current study focused on psychodynamic therapy, it is possible that SC is a curative factor common across many other schools of psychotherapy. Future studies should examine SC in the context of other forms of treatments.

Our results also extend previous findings by demonstrating the within-client association between SC and session-level outcomes. This significant association held even when we controlled for the outcome variable levels in the previous session; this approach reduces concerns regarding reverse causation and emphasizes the possible role of SC as a process variable that contributes to treatment outcomes. Of course, the relationship between SC and outcome could be more complex, and reverse causality could also occur. Future studies should examine the reciprocal effects of changes in symptoms and functioning level on clients' SC. Moreover, while these questions are beyond the scope of the current study, future studies should explore specific patterns of change in SC and their effect on treatment outcome.

4.1 | Limitations

Several limitations of this study should be noted. First, this study occurred in a university community clinic where the therapists are trainees. Although all therapists received intensive supervision, their status may limit our ability to generalize the results to more experienced clinicians. Furthermore, because a substantial group of therapists in our sample treated only one client each, we could not estimate therapist effects (Schiefele et al., 2017). Future studies with larger numbers of clients per therapist are needed to examine the effect of specific therapists' characteristics on the development of SC during treatment and its association with treatment outcome.

Second, our sample included clients seeking treatment for a variety of problems and diagnoses. However, some of our clients did not meet criteria for a *DSM-IV* diagnosis. More research is needed to generalize our results to additional populations, such as clients diagnosed with a specific *DSM* condition or to a clinical population with more severe conditions.

Third, the study was designed as a naturalistic field study and did not include a comparison or control group. Because the therapeutic interventions were not administered using a specific protocol, we did not conduct adherence tests; hence, the results cannot be linked exclusively to the psychotherapy treatment offered or to a specific therapeutic orientation. To shed light on this matter, future studies should explore the association between changes in clients' SC levels and specific therapeutic interventions.

Fourth, clients' weekly levels of SC were measured using a brief (three-item) adaptation of the full SCS (Neff, 2003b). These items were selected based on their correlation with the general scale as well as with the individual subscales (Raes et al., 2011). However, more research is needed to establish the validity and reliability of this adapted brief scale. The fact that clients were asked to report how self-compassionate they were on a weekly basis might have had an effect on their level of SC or on the focus of the therapy. In addition, some caution should be taken in interpreting the results because both our session-level and treatment-level measures were self-reported. Future studies should consider utilizing explicit or objective methods to measure SC and outcome.

Finally, the statistical analyses conducted in this study were correlational, and thus our findings could be attributed to an unmeasured, uncontrolled third variable that may account for the observed changes in both symptoms and SC. We used lagged scores for both the SC and outcome measures to reduce the threat of reverse causation. Nevertheless, our results are correlational; therefore, causality cannot be explicitly assumed.

5 | CONCLUSION

The results of the current study highlight SC as a promising target for psychotherapeutic interventions among clients suffering from a variety of emotional and psychological conditions. The findings accord with previous studies (e.g., Blatt et al., 1995; Hawley et al., 2006) in recognizing the problematic outcomes of clients' self-criticism and rumination. Interestingly, our session-level SC measure, which focused on the positive aspects of SC (vs. the negative aspects such as self-criticism), predicted clients' level of functioning and symptomology. These results suggest that the positive aspects of SC were strongly associated with positive outcomes.

One clinical implication of this study is that clinicians should aim to help their clients develop the positive qualities of SC. The therapeutic relationship should serve as a collaborative setting in which clients can learn to create a kind and supportive inner dialogue and accept their difficulties—and even their failures—as part of a shared human experience. One way to do so is through greater integration of mindfulness or acceptance theory (and practice) into the more traditional, psychodynamic therapeutic framework. Such integration may in turn decrease symptoms of depression and anxiety and contribute to improved treatment outcomes.

NOTE

¹ The following DSM-IV diagnoses were assumed in the affective disorders cluster: major depressive disorder, dysthymia and bipolar disorder. The following DSM-IV diagnoses were assumed in the anxiety disorders cluster: panic disorder, agoraphobia, generalized anxiety disorder and social anxiety disorder.

² To address missing posttreatment data, we conducted multivariate imputations using chained equations (Azur, Stuart, Frangakis, & Leaf, 2011), as implemented in SAS PROC MI and PROC Analyze. The results from the imputed data were very similar to those from the available data (with the exception of the OQ-45 analysis that showed a similar trend but did not reach significance) and therefore will not be reported further.

³ A similar pattern of result is found when the therapeutic alliance is not included as a covariate in the model.

⁴ For HSCL, the level 3 variance accounted for 6% of the total variance but was not significant. We re-ran the model as a 3-level MLM with very similar results.

ORCID

Lior Galili-Weinstock  <http://orcid.org/0000-0003-3581-4066>

REFERENCES

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Aronfreed, J. J. (1964). The origin of self-criticism. *Psychological Review*, 71(3), 193–218.
- Barnard, L. K., & Curry, J. F. (2011). Self-compassion: Conceptualizations, correlates, & interventions. *Review of General Psychology*, 15(4), 289–303. <https://doi.org/10.1037/a0025754>
- Beck, A., Steer, R., & Brown, G. (1996). BDI-II. Beck Depression Inventory-second edition. Retrieved from <https://www.nctsn.org/content/beck-depression-inventory-second-edition-bdi-ii>
- Blagys, M. D., & Hilsenroth, M. J. (2000). Distinctive feature of short-term psychodynamic-interpersonal psychotherapy: A review of the comparative psychotherapy process literature. *Clinical Psychology: Science and Practice*, 7, 167–188. <https://doi.org/10.1093/clipsy.7.2.167>
- Blatt, S. J. (1974). Levels of object representation in anaclitic and introjective depression. *The Psychoanalytic Study of the Child*, 29, 107–157. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/4445397>
- Blatt, S. J. (1995). The destructiveness of perfectionism. *Implications for the treatment of depression*. *The American Psychologist*, 50(12), 1003–1020. <https://doi.org/10.1037/0003-066X.50.12.1003>
- Blatt, S. J., Quinlan, D. M., Chevron, E. S., McDonald, C., & Zuroff, D. (1982). Dependency and self-criticism: Psychological dimensions of depression. *Journal of Consulting and Clinical Psychology*, 50(1), 113–124. <https://doi.org/10.1037/0022-006X.50.1.113>
- Blatt, S. J., Quinlan, D. M., Pilkonis, P. A., & Shea, M. T. (1995). Impact of perfectionism and need for approval on the brief treatment of depression: The National Institute of Mental Health Treatment of Depression Collaborative

- Research Program revisited. *Journal of Consulting and Clinical Psychology*, 63(1), 125–132. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/7896977>
- Blatt, S. J., Zuroff, D. C., Bondi, C. M., Sanislow, C. A., & Pilkonis, P. A. (1998). When and how perfectionism impedes the brief treatment of depression: Further analyses of the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology*, 66(2), 423–428.
- Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research and Practice*, 16, 252–260. Retrieved from <https://doi.org/10.1037/h0085885>
- Brach, T. (2004). Radical acceptance: Embracing your life with the heart of a Buddha. Retrieved from <https://www.amazon.co.uk/Radical-Acceptance-Embracing-Heart-Buddha/dp/1452606021>
- Callahan, J. L., Aubuchon-Endsley, N., Borja, S. E., & Swift, J. K. (2009). Pretreatment expectancies and premature termination in a training clinic environment. *Training and Education in Professional Psychology*, 3(2), 111–119. <https://doi.org/10.1037/a0012901>
- Castilho, P., Pinto-Gouveia, J., & Duarte, J. (2015). Evaluating the multifactor structure of the long and short versions of the self-compassion scale in a clinical sample. *Journal of Clinical Psychology*, 0(0), 1–15. <https://doi.org/10.1002/jclp.22187>
- Costa, J., Marôco, J., Pinto-Gouveia, J., Ferreira, C., & Castilho, P. (2015). Validation of the psychometric properties of the Self-Compassion Scale. Testing the Factorial Validity and Factorial Invariance of the Measure among Borderline Personality Disorder, Anxiety Disorder, Eating Disorder and General Populations. *Clinical Psychology & Psychotherapy*. <https://doi.org/10.1002/cpp.1974>
- Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably? *Personality and Social Psychology Bulletin*, 32(7), 917–929. <https://doi.org/10.1177/0146167206287721>
- Crits-Christoph, P., Connolly Gibbons, M., & Mukherjee, D. (2013). Process outcome research. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavioral change* (6th ed., pp. 298–340). New York: Wiley.
- Dalai Lama (2001). *An open heart: Practicing compassion in everyday life*. Boston. Retrieved from <https://www.amazon.com/Open-Heart-Practicing-Compassion-Everyday/dp/0316989797>
- Derogatis, C. R. (1977). HSCL-90, administration, scoring, and procedures. Manual 1 for the R(evised) version and other instruments of the Psychopathology Rating Scale Series. Baltimore: John Hopkins University School of Medicine.
- Diedrich, A., Grant, M., Hofmann, S. G., Hiller, W., & Berking, M. (2014). Self-compassion as an emotion regulation strategy in major depressive disorder. *Behaviour Research and Therapy*, 58, 43–51. <https://doi.org/10.1016/j.brat.2014.05.006>
- Ferreira, C., Matos, M., Duarte, C., & Pinto-gouveia, J. (2014). Shame memories and eating psychopathology: The buffering effect of self-compassion. *European Eating Disorders Review*, 22, 487–494. <https://doi.org/10.1002/erv.2322>
- Finlay-Jones, A. L., Rees, C. S., & Kane, R. T. (2015). Self-Compassion, emotion regulation and stress among Australian psychologists: Testing an emotion regulation model of self-compassion using structural equation modeling. *PLoS ONE*, 10(7), 1–19. <https://doi.org/10.1371/journal.pone.0133481>
- Fox, H. C., Hong, K. A., & Sinha, R. (2008). Difficulties in emotion regulation and impulse control in recently abstinent alcoholics compared with social drinkers. *Addictive Behaviors*, 33(2), 388–394. <https://doi.org/10.1016/j.addbeh.2007.10.002>
- Germer, C. K., & Neff, K. (2013). Self-compassion in clinical practice. *Journal of Clinical Psychology*, 69(8), 856–867. <https://doi.org/10.1002/jclp.22021>
- Gilbert, P. (2000). Social mentalities: Internal 'social' conflict and the role of inner warmth and compassion in cognitive therapy. *In Genes on the couch: Explorations in evolutionary psychotherapy*. (pp. 118–150). New York: Brunner-Routledge.
- Gilbert, P. (2010). An introduction to the theory & practice of compassion focused therapy and compassionate mind training for shame based difficulties. Retrieved from <http://www.communityofmindfulparenting.com/curriculum/rrdocuments/rrCompassionateMindTrainig.pdf>
- Gilbert, P., & Procter, S. (2006). For people with high shame and self-criticism: Overview and pilot study of a group therapy approach. *Clinical Psychology & Psychotherapy*, 13, 353–379. <https://doi.org/10.1002/cpp.507>
- Gratz, K. L., & Gunderson, J. G. (2006). Preliminary data on an acceptance-based emotion regulation group intervention for deliberate self-harm among women with borderline personality disorder. *Behavior Therapy*, 37(1), 25–35. <https://doi.org/10.1016/j.beth.2005.03.002>
- Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation. *Journal of Psychopathology and Behavioral Assessment*, 26(1), 41–54. <https://doi.org/10.1023/B:JOBA.0000007455.08539.94>
- Greenberg, L. S., & Watson, J. C. (2006). *Emotion-focused therapy for depression*. Washington DC: American Psychological Association.

- Hatcher, R. L., & Gillaspay, J. A. (2006). Development and validation of a revised short version of the working alliance inventory. *Psychotherapy Research*, 16(August 2011), 12–25. <https://doi.org/10.1080/10503300500352500>
- Hawley, L. L., Ho, M.-H. R., Zuroff, D. C., & Blatt, S. J. (2006). The relationship of perfectionism, depression, and therapeutic alliance during treatment for depression: Latent difference score analysis. *Journal of Consulting and Clinical Psychology*, 74(5), 930–942. <https://doi.org/10.1037/0022-006X.74.5.930>
- Hayes, J. A., Lockard, A. J., Janis, R. A., & Locke, B. D. (2016). Construct validity of the Self-Compassion Scale-Short Form among psychotherapy clients. *Counselling Psychology Quarterly*, 5070(May), 0–18. <https://doi.org/10.1080/09515070.2016.1138397>
- Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2002). Review of “Acceptance and commitment therapy: An experimental approach to behavior change.” *Child Family Behavior Therapy*, 24(4), 51–57. https://doi.org/10.1300/J019v24n04_04
- Heffernan, M., Quinn Griffin, M. T., McNulty, S. R., & Fitzpatrick, J. J. (2010). Self-compassion and emotional intelligence in nurses. *International Journal of Nursing Practice*, 16, 366–373. <https://doi.org/10.1111/j.1440-172X.2010.01853.x>
- Hollis-Walker, L., & Colosimo, K. (2011). Mindfulness, self-compassion, and happiness in non-meditators: A theoretical and empirical examination. *Personality and Individual Differences*, 50, 222–227. <https://doi.org/10.1016/j.paid.2010.09.033>
- Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy*, 48(1), 9–16. <https://doi.org/10.1037/a0022186>
- Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, 36(2), 223–233. <https://doi.org/10.1037/0022-0167.36.2.223>
- Keijsers, G. P. J., Kampman, M., & Hoogduin, C. A. L. (2001). Dropout prediction in cognitive behavior therapy for panic disorder. *Behavior Therapy*, 32(4), 739–749. [https://doi.org/10.1016/S0005-7894\(01\)80018-6](https://doi.org/10.1016/S0005-7894(01)80018-6)
- Krieger, T., Altenstein, D., Baettig, I., Doerig, N., & Holtforth, M. G. (2013). Self-compassion in depression: Associations with depressive symptoms, rumination, and avoidance in depressed outpatients. *Behavior Therapy*, 44(3), 501–513. <https://doi.org/10.1016/j.beth.2013.04.004>
- Kuyken, W., Watkins, E., Holden, E., White, K., Taylor, R. S., Byford, S., ... Dalglish, T. (2010). How does mindfulness-based cognitive therapy work? *Behaviour Research and Therapy*, 48(11), 1105–1112. <https://doi.org/10.1016/j.brat.2010.08.003>
- Lambert, M. J., Burlingame, G. M., Umphress, V., Hansen, N. B., Vermeersch, D. A., Clouse, G. C., & Yanchar, S. C. (1996). The reliability and validity of the Outcome Questionnaire. *Clinical Psychology & Psychotherapy*, 3(4), 249–258. [https://doi.org/10.1002/\(SICI\)1099-0879\(199612\)3:4<249::AID-CPP106>3.0.CO;2-5](https://doi.org/10.1002/(SICI)1099-0879(199612)3:4<249::AID-CPP106>3.0.CO;2-5)
- Lockard, A. J., Hayes, J. A., Neff, K., & Locke, B. D. (2014). Self-compassion among college counseling center clients: An examination of clinical norms and group differences. *Journal of College Counseling*, 17(3), 249–259. <https://doi.org/10.1002/j.2161-1882.2014.00061.x>
- Lutz, W., Leon, S. C., Martinovich, Z., Lyons, J. S., & Stiles, W. B. (2007). Therapist effects in outpatient psychotherapy: A three-level growth curve approach. *Journal of Counseling Psychology*, 54(1), 32–39. <https://doi.org/10.1037/0022-0167.54.1.32>
- Lutz, W., Tholen, S., Schürch, E., & Berking, M. (2006). Die Entwicklung, Validierung und Reliabilität von Kurzformen Gängiger Psychometrischer Instrumente zur Evaluation des Therapeutischen Fortschritts in Psychotherapie und Psychiatrie. [The development, validation, and reliability of short-forms of current instruments for the evaluation of therapeutic progress in psychotherapy and psychiatry] *Diagnostica*, 52, 11–25. <https://doi.org/10.1026/0012-1924.52.1.11>
- MacBeth, A., & Gumley, A. (2012). Exploring compassion: A meta-analysis of the association between self-compassion and psychopathology. *Clinical Psychology Review*, 32(6), 545–552. <https://doi.org/10.1016/j.cpr.2012.06.003>
- Magnus, C. M. R., Kowalski, K. C., & McHugh, T.-L. F. (2010). The role of self-compassion in women's self-determined motives to exercise and exercise-related outcomes. *Self and Identity*, 9(February 2015), 363–382. <https://doi.org/10.1080/15298860903135073>
- Miller, S. D., Duncan, B. L., Brown, J., Sparks, J. A., & Claud, D. A. (2003). The Outcome Rating Scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy*, 2, 2, 91–100.
- Neacsiu, A. D., Eberle, J. W., Kramer, R., Wiesmann, T., & Linehan, M. M. (2014). Dialectical behavior therapy skills for transdiagnostic emotion dysregulation: A pilot randomized controlled trial. *Behavioral Research and Therapy*, 59, 40–51. <https://doi.org/10.1016/j.brat.2014.05.005>
- Neff, K. (2003a). Self-compassion: An alternative conceptualization of a healthy attitude toward oneself. *Self and Identity*, 2(2), 85–101. <https://doi.org/10.1080/15298860309032>
- Neff, K. (2003b). The development and validation of a scale to measure self-compassion. *Self and Identity*, 2(3), 223–250. <https://doi.org/10.1080/15298860309027>
- Neff, K., & Germer, C. (2013). A pilot study and randomized controlled trial of the mindful self-compassion program. *Journal of Clinical Psychology*, 69(1), 28–44. <https://doi.org/10.1002/jclp.21923>

- Neff, K., Hsieh, Y.-P., & Dejjtterat, K. (2005). Self-compassion, Achievement Goals, and Coping with Academic Failure. *Self and Identity*, 4(3), 263–287. <https://doi.org/10.1080/13576500444000317>
- Neff, K., Kirkpatrick, K. L., & Rude, S. S. (2007). Self-compassion and adaptive psychological functioning. *Journal of Research in Personality*, 41(1), 139–154. <https://doi.org/10.1016/j.jrp.2006.03.004>
- Neff, K. D. (2016). The Self-Compassion Scale is a valid and theoretically coherent measure of self-compassion. *Mindfulness*, 7, 264–274. <https://doi.org/10.1007/s12671-015-0479-3>
- Neff, K. D., Whittaker, T. A., & Karl, A. (2017). Examining the factor structure of the Self-Compassion Scale in four distinct populations: Is the use of a total scale score justified? *Journal of Personality Assessment*, 1–12. <https://doi.org/10.1080/00223891.2016.1269334>
- Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology*, 48(1), 85–112. <https://doi.org/10.1016/j.jsp.2009.09.002>
- Raes, F., Pommier, E., Neff, K., & Van Gucht, D. (2011). Construction and factorial validation of a short form of the Self-Compassion Scale. *Clinical Psychology & Psychotherapy*, 18(3), 250–255. <https://doi.org/10.1002/cpp.702>
- Raudenbush, S., & Byrk, A. (2001). Hierarchical linear models: Applications and data analysis methods (advanced quantitative techniques in the social sciences). Retrieved from <https://www.amazon.com/Hierarchical-Linear-Models-Applications-Quantitative/dp/076191904X>
- Roemer, L., Lee, J. K., Salters-Pedneault, K., Erisman, S. M., Orsillo, S. M., & Mennin, D. S. (2009). Mindfulness and emotion regulation difficulties in generalized anxiety disorder: Preliminary evidence for independent and overlapping contributions. *Behavior Therapy*, 40(2), 142–154. <https://doi.org/10.1016/j.beth.2008.04.001>
- Salzberg, S. (1997). A heart as wide as the world: Living with mindfulness, wisdom, and compassion. Retrieved from https://www.google.co.il/books?hl=iw&lr=&id=2NiAn_6sEjC&pgis=1
- Sbarra, D. A., Smith, H. L., & Mehl, M. R. (2012). When leaving your ex, love yourself: Observational ratings of self-compassion predict the course of emotional recovery following marital separation. *Psychological Science*, 23, 261–269. <https://doi.org/10.1177/0956797611429466>
- Schanche, E., Stiles, T. C., McCullough, L., Svartberg, M., & Nielsen, G. H. (2011). The relationship between activating affects, inhibitory affects, and self-compassion in patients with Cluster C personality disorders. *Psychotherapy*, (Chicago, Ill.), 48(3), 293–303. <https://doi.org/10.1037/a0022012>
- Scharff, J. S., & Tsigounis, S. A. (2003). *Self-hatred in psychoanalysis: Detoxifying the persecutory object*. New York: Routledge.
- Schiefele, A.-K., Lutz, W., Barkham, M., Rubel, J., Böhnke, J., Delgado, J., ... Lambert, M. J. (2017). Reliability of therapist effects in practice-based psychotherapy research: A guide for the planning of future studies. *Administration and Policy in Mental Health and Mental Health Services Research*, 44(5), 598–613. <https://doi.org/10.1007/s10488-016-0736-3>
- Shahar, G. (2001). Personality, shame, and the breakdown of social bonds: The voice of quantitative depression research. *Psychiatry*, 64(3), 228–239. <https://doi.org/10.1521/psyc.64.3.228.18463>
- Shedler, J. (2010). The efficacy of psychodynamic psychotherapy. *The American Psychologist*, 65(2), 98–109. <https://doi.org/10.1037/a0018378>
- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., ... Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, 59, 22–33. [https://doi.org/10.1016/S0924-9338\(99\)80239-9](https://doi.org/10.1016/S0924-9338(99)80239-9)
- Shrout, P., & Lane, S. P. (2012). Psychometrics. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 301–320). New York: Guilford Press.
- Singer, J. D., & Willett, J. B. (2003). Applied longitudinal data analysis: Modeling change and event occurrence. Retrieved from <https://books.google.com/books?hl=iw&lr=&id=PpnA1M8VwR8C&pgis=1>
- Storch, E. A., Roberti, J. W., & Roth, D. A. (2004). Factor structure, concurrent validity, and internal consistency of the Beck Depression Inventory-Second Edition in a sample of college students. *Depression and Anxiety*, 19(3), 187–189. <https://doi.org/10.1002/da.20002>
- Teasdale, J. D., Segal, Z. V., Williams, J. M., Ridgeway, V. A., Soulsby, J. M., & Lau, M. A. (2000). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *Journal of Consulting and Clinical Psychology*, 68(4), 615–623. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/10965637>
- Vettese, L. C., Dyer, C. E., Li, W. L., & Wekerle, C. (2011). Does self-compassion mitigate the association between childhood maltreatment and later emotion regulation difficulties? A preliminary investigation. *International Journal of Mental Health and Addiction*, 9(5), 480–491. <https://doi.org/10.1007/s11469-011-9340-7>
- Werner, K., Jazaieri, H., Goldin, P., Ziv, M., Hiemberg, R., & Gross, J. (2012). Self compassion and social anxiety disorder. *Anxiety, Stress and Coping*, 25(5), 543–558.

- Westmacott, R., Hunsley, J., Best, M., Rumstein-McKean, O., & Schindler, D. (2010). Client and therapist views of contextual factors related to termination from psychotherapy: A comparison between unilateral and mutual terminators. *Psychotherapy Research, 20*(4), 423–435. <https://doi.org/10.1080/10503301003645796>
- Young, J., Klosko, E., & Wieshar, M. (2003). *Schema therapy: A practitioner's guide*. New York: Guilford Press. <https://doi.org/10.1176/appi.ajp.160.11.2074-a>
- Zessin, U., Dickhäuser, O., & Garbade, S. (2015). The relationship between self-compassion and well-being: A meta-analysis. *Applied Psychology: Health and Well-Being, 7*(3), 340–364. <https://doi.org/10.1111/aphw.12051>

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Galili-Weinstock L, Chen R, Atzil-Slonim D, Bar-Kalifa E, Peri T, Rafaeli E. The association between self-compassion and treatment outcomes: Session-level and treatment-level effects. *J Clin Psychol.* 2017;00:1–18. <https://doi.org/10.1002/jclp.22569>