Psychotherapy Research

Narrative Therapy vs. Cognitive-Behavioral Therapy for moderate depression: Empirical evidence from a controlled clinical trial

Rodrigo T. Lopes[a], Miguel M. Gonçalves[a], Paulo P.P. Machado[a], Dana Sinai[b], Tiago Bento[c] & João Salgado[c]

[a] School of Psychology, University of Minho, Braga, Portugal

[b] Department of Psychology, Ben Gurion University of the Negev, Be'er Sheva, Israel

[c] Instituto Superior da Maia, CINEICC/ISMAI, Maia, Portugal
Published online: 30 Jan 2014.

PLEASE SCROLL DOWN FOR ARTICLE

Routledge
Taylor & Francis Group

EMPIRICAL PAPER

Narrative Therapy vs. Cognitive-Behavioral Therapy for moderate depression: Empirical evidence from a controlled clinical trial

RODRIGO T. LOPES[1], MIGUEL M. GONÇALVES[1], PAULO P.P. MACHADO[1], DANA SINAI[2], TIAGO BENTO[3], & JOÃO SALGADO[3]

[1]*School of Psychology, University of Minho, Braga, Portugal;* [2]*Department of Psychology, Ben Gurion University of the Negev, Be'er Sheva, Israel &* [3]*Instituto Superior da Maia, CINEICC/ISMAI, Maia, Portugal*

Abstract
Background: Systematic studies of the efficacy of Narrative Therapy (NT) for depression are sparse. **Objective**: To evaluate the efficacy of individual NT for moderate depression in adults compared to Cognitive-Behavioral Therapy (CBT). **Method**: Sixty-three depressed clients were assigned to either NT or CBT. The Beck Depression Inventory-II (BDI-II) and Outcome Questionnaire-45.2 (OQ-45.2) were used as outcome measures. **Results**: We found a significant symptomatic reduction in both treatments. Group differences favoring CBT were found on the BDI-II, but not on the OQ-45.2. **Conclusions**: Pre- to post-treatment effect sizes for completers in both groups were superior to benchmarked waiting-list control groups.

Keywords: depression; treatment of depression; psychological treatment of depression; Empirically Supported Therapy (EST); Narrative Therapy; Cognitive-Behavioral Therapy

Depression is among the most common mental health problems for which help is sought (Andrews & Thomson, 2009; Roness, Mykletun, & Dahl, 2005). Despite the fact that 12 treatments for unipolar depression are currently listed as having reliable empirical support (for a complete list, see Hayes and Strunk, n.d.), a considerable proportion of clients leave treatment without substantial gains and/or relapse after treatment. Thus, outcome research with an emphasis on treatment efficacy is recommended (Chambless & Hollon, 1998; Chambless et al., 1998; Kendall, 1998; Kendall, Holmbeck, & Verduin, 2004).

Narrative Therapy (NT; White, 2007; White & Epston, 1990) is a psychotherapeutic approach based on the notion that people construct narratives to define themselves and give meaning to their daily experiences and life events. Psychological suffering is viewed as a problem-saturated way of constructing life stories and a person's identity. These rigid self-narratives constrain the person's actions, feelings, and thoughts (White, 2007) and obscure life alternatives. The purpose of psychotherapy is to help clients narrate their life stories in richer and more gratifying ways. This view of the human being as a meaning-making agent has had a considerable impact on clinical psychology and psychotherapy in recent years (Angus & McLeod, 2004).

Although widely practiced throughout the world, NT remains underresearched (Busch, 2011; Chenail, DeVicentis, Kiviat, & Somers, 2012; Etchison & Kleist, 2000). Some qualitative process research studies (Matos, Santos, Gonçalves, & Martins, 2009; Moreira, Beutler, & Gonçalves, 2008) and several case studies (Betchley & Falconer, 2002; Cashin, 2008; da Costa, Nelson, Rudes, & Guterman, 2007; Draucker, 1998; Kropf & Tandy, 1998; Nylund, 2002; Palgi & Ben-Ezra, 2010; Rothschild, Brownlee, & Gallant, 2000; Young, 2008) suggest positive outcomes for a wide range of problems and disorders. However, only one of these qualitative studies

---

Correspondence concerning this article should be addressed to Miguel M. Gonçalves, University of Minho, School of Psychology, Braga, Portugal. Email: mgoncalves@psi.uminho.pt

(Kropf & Tandy, 1998) has addressed the effectiveness of NT for depression, and none have done so in a controlled clinical trial. To the best of our knowledge, only one study tested the efficacy of NT (White & Epston, 1990) in a large sample of adults with moderate depression using a treatment manual. Vromans and Schweitzer (2011) found a large pre- to post-treatment effect size ($d = 1.36$) for those who completed the NT intervention (38 out of the initial 47 clients). Moreover, a clinical significance analysis performed with completers according to Jacobson and Truax's (1991) criteria indicated that 74% of the participants improved reliably, 61% moved to the functional population range, and 53% recovered. It was also found that NT completers showed significant improvements on a measure of interpersonal difficulties. Intention-to-treat (ITT) analysis obtained a large pre-post effect size ($d = 1.10$) that was comparable to other treatments. Although this study had methodological strengths, such as the use of a structured interview for diagnostic assessment at pre-treatment, assessment of clinically significant change, and follow-up assessment at 3 months, it lacked a comparison group.

Because establishing a waiting list control group or administering a placebo in clinical trials is often ethically controversial, Chambless and Hollon (1998) suggested comparison with an empirically supported intervention for the same population. Cognitive-Behavioral Therapy for depression (CBT; Beck, Rush, Shaw, & Emery, 1979) is a strong candidate for a comparison group in the treatment of depression. It has received consistent empirical support since the first study in the 1970s (Rush, Beck, Kovacs, & Hollon, 1977), and it remains one of the best-known psychological treatments for depression (Butler, Chapman, Forman, & Beck, 2006; Cuijpers, van Straten, Andersson, & van Oppen, 2008; DeRubeis & Crits-Christoph, 1998; Dobson, 1989; Gloaguen, Cottraux, Cucherat, & Blackburn, 1998).

The present study aims to test the efficacy of NT in a comparative controlled trial with CBT as the comparison group and a primary focus on the reduction of depressive symptoms. The secondary aim of this study is to assess the reduction of general psychological distress. Thus, the specific research questions were (1) whether NT is as efficacious as CBT in reducing depressive symptoms in adults; (2) whether clients in the two treatments experience changes at different rates; and (3) whether the dropout rates differ significantly between the two treatments. It was expected that NT outcomes would be comparable to CBT.

Methods

Participants

Clients. Eligibility criteria were (a) having a diagnosis of Major Depressive Disorder according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; American Psychiatric Association, 2000), (b) being over 18 years old, (c) agreeing to sign an informed consent to treatment, and (d) agreeing to the videotaping of sessions and the completion of research questionnaires. Participants were excluded if they presented (a) any Axis II diagnosis, (b) any other concurrent Axis I disorder that might be a focus of clinical attention (substance-related disorders, sexual disorders, eating disorders), (c) severe suicidal ideation, (d) psychotic symptoms, or (e) bipolar disorder. Clients with anxiety complaints or a secondary anxiety disorder were not excluded if the anxiety was not considered a primary complaint (e.g., panic disorder).

The trial was conducted at a psychology university clinic in the north of Portugal that offered treatment to both the on- and off-campus community. Potential participants were either self-referred or referred by community mental health professionals. Therapy was offered free of charge.

When potential participants sought consultation, they were referred to the screening staff to determine study eligibility. Eligible participants were assigned by the clinic secretary to a treatment condition according to their incoming order—one to NT, the next to CBT, and so on. The secretary was unaware of the specific procedures of each treatment. The intake form was given to the therapist, who made contact with the client for a first appointment. The participants were unaware of the treatment modality they would receive.

Therapists. To describe the level of clinical experience, we used the definition proposed by Orlinsky and Rønnestad (2005). The study included six novice therapists (less than 1.5 years of clinical experience), one apprentice therapist (between 1.5 and 3.5 years of clinical experience), and three graduate therapists (between 3.5 and 7 years of clinical experience). Two seasoned therapists (between 15 and 25 years of clinical experience) supervised the study's preparation and implementation, from the elaboration of the NT manual and adaptation of the CBT manual to training and supervision. Some therapists delivered CBT and others delivered NT in a nested design (i.e., they only treated clients in one treatment manual). One supervisor was responsible for the NT group (second author), and the other supervisor was responsible for the CBT group (sixth author). Both had advanced

specialized training and extensive clinical and teaching experience with the psychotherapy approach they supervised.

The three graduate therapists (years of experience $_{Mean}$ = 5; SD = 0.57) treated 70.8% of the cases, the apprentice therapist (3 years of clinical experience) treated 3.1% of the cases, and the six novice therapists (years of experience$_{Mean}$ = .33; SD = 0.52) treated 26.2% of the cases. The therapists had an average of 1.9 years of experience (SD = 2.13). The proportion of cases treated by novice therapists was significantly larger in the CBT group compared to the NT group (χ^2 (2) = 13.071, p < .002).

All of the therapists were psychologists and graduate students of the Clinical Psychology Program (six at the master's level and four at the PhD level) at the time the trial took place. In Portugal, psychologists complete a 5-year program (master's degree), including a 1-year clinical internship, which is the minimum training required for supervised practice. Reported years of experience do not include the internship.

Therapists' training. One of the graduate therapists (first author) had previous experience conducting CBT. The other graduate therapist (fifth author) collaborated in the development of the NT treatment manual and had previous experience conducting NT. The other graduate therapist and the apprentice therapist had both theoretical and practical training in NT before joining the project. They all received specific training on the respective treatment manuals from the seasoned therapists. The six novice therapists had theoretical training before joining the project, and they received specific training in the treatment manuals, which consisted of watching videotapes of experienced NT or CBT psychotherapists with depressed clients. They also received close supervision by the two graduate therapists and the two seasoned therapists.

Procedures

Measures. *Structured Clinical Interview for DSM-IV - Axis I Disorders, Clinician version* (SCID-I; First, Spitzer, Gibbons, & Williams, 2002) and *Structured Clinical Interview for DSM-IV - Axis II Personality Disorders* (SCID-II; First, Gibbon, Spitzer, Williams, & Benjamin, 1997). The SCID-I and SCID-II are the most frequently used diagnostic interviews in research. Training with these clinical interviews consisted of novice therapists watching graduate therapists perform the clinical interviews. These graduate therapists later assisted the novice therapists in their first interviews. The diagnostic

decision was discussed in small groups of therapists coordinated by a graduate therapist.

The Graffar Index. Socioeconomic status (SES) was measured with the Graffar index (Graffar, 1956), which is a short scale that takes into account income, level of education, profession, and type of home. It divides the population into five socioeconomic layers. The data are shown using a continuous score ranging from 5 (highest SES) to 25 (lowest SES).

Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996). The BDI-II is considered the gold standard measure for all research on depression and was the primary outcome measure in this study. It consists of 21 self-reported items grouped into three subscales: Cognitive symptoms, affective symptoms, and somatic symptoms. The total score ranges from 0 to 63. The BDI-II has been shown to have high internal consistency (α = 0.91; Steer, Brown, Beck, & Sanderson, 2001; and α = 0.89 in our sample). It has been translated and validated for the Portuguese population with data similar to the American sample (Campos & Gonçalves, 2011; Coelho, Martins, & Barros, 2002). Because the Portuguese studies did not calculate the Reliable Change Index (RCI; Jacobson & Truax, 1991), normative data gathered from meta-analyses of diverse samples (Seggar, Lambert, & Hansen, 2002) were used to calculate the proportion of clinical change (RCI = 8.46; cut-off score = 14.29).

Outcome Questionnaire (OQ-45.2; Lambert et al., 1996). The OQ-45.2 is a self-report questionnaire designed to assess clients' progress throughout therapy and its termination. It comprises 45 questions concerning psychological distress and interpersonal relations as well as the social role of the client. Cronbach's alpha was .93 in various clinical samples (de Jong et al., 2007; Lambert et al., 1996), and similar reliability was obtained in our sample (α = 0.89), which is indicative of excellent internal consistency. The RCI calculated for the Portuguese population was 15 points, and the cutoff was 62 points (Machado & Fassnacht, 2014).

Adherence and Competence Scale for Narrative and Cognitive-Behavioral Therapy (ACS-N-CBT; Gonçalves, Bento, Lopes, & Salgado, 2009). This scale was developed specifically for this trial and allows the quality of the sessions to be rated from the videotapes. It consists of five "yes or no" questions concerning general therapeutic attitudes (therapist's empathy, warmth, ability to control the time of session) and 10 "never to always" questions

(0 to 6 Likert-type scale) concerning the frequency with which the therapist made use of the therapeutic techniques specific to NT and CBT in the particular session. Raters were also asked to use a Likert scale (from 0 to 6) to evaluate the therapists' competence in using each of the specific aforementioned techniques. Thus, this scale yields five scores: (1) the use of general therapeutic attitudes in the session; (2) the frequency with which CBT techniques were used in the session; (3) the frequency with which NT techniques were used in the session; (4) general psychotherapeutic competence; and (5) competence in using specific therapeutic techniques from CBT and/or NT. The scale developers provided training in this scale with examples of random sessions. Care was taken that the chosen example sessions were not the same as the ones assigned to rating. A manual for the ACS-N-CBT scale containing instructions for each item of the scale and common questions was provided to the raters as support during the rating process.

Assessment procedures. The SCID-I and SCID-II interviews were used in the initial assessment to determine eligibility and were conducted by therapists trained for this purpose. No inter-rater reliability was calculated because one evaluator conducted each interview. Questions about diagnostic criteria were discussed with the supervisors. The SCID-I (First et al., 2002) was also used to gather demographic and clinical information about the participants. The demographic variables assessed were age, gender, relationship status, professional status, and education. The clinical variables assessed were the presence of co-morbidity, medication use at intake, previous hospitalizations, previous suicide attempts, and previous experiences with psychotherapy. To assess the social, occupational, and psychological functioning of the participants, the Global Assessment of Functioning (GAF) scale was used. The BDI-II and OQ-45.2 were used to measure the baseline severity of depressive symptoms and treatment outcomes. BDI-II was defined as the primary outcome measure, and OQ-45.2 was defined as the secondary outcome measure. Questionnaires were given to the participants before the initial session and at every fourth session (i.e., at sessions 1, 4, 8, 12, 16, and 20). The assessment of treatment integrity using the ACS-N-CBT was made after the trial was completed.

Treatment conditions. Treatment in both conditions consisted of 20 sessions of 1 hr each. Sessions took place weekly from session 1 to 16 and every 2 weeks from session 17 to 20. Termination occurred at the end of the 20 sessions, as scheduled by the treatment manual (regardless of the clinical status of the client), or at an earlier stage of the treatment manual that was mutually agreed upon between the therapist and client (in cases where the client had reached the therapeutic goals).

Narrative Therapy. NT focuses on the role of narrative processes in the organization of experience, knowledge, and behavior. Problems arise as autobiographical narratives and are restricted to problematic contents, which White and Epston (1990) call "problem saturated narratives." The therapist engages the client in activities to "re-write" self-narratives in a process organized around three phases: A deconstruction phase, a reconstruction phase, and a consolidation and termination phase (Freedman & Combs, 1996). During the deconstruction phase, the central goal is to understand the problem in its context, circumstances, assumptions, effects, and influence on the person's life in an attempt to separate the problem from the person (mainly using externalization). In the reconstruction phase, the main goal is to expand the narrative elaboration of novelties, or experiences that fall outside the domain of the problem-saturated narrative, gradually facilitating the emergence of an alternative self-narrative. These new narrative elements, called unique outcomes (White, 2007), are explored and contribute to an alternative self-narrative that redefines people's relationship with themselves, their history, and their significant others. In the last phase of the protocol, referred to as the consolidation and termination phase, the goal is to consolidate the alternative self-narrative and to root it in the network of socio-cultural discourses and practices.

Following previous criticism on the rigidity of therapeutic manuals (Ablon, Levy, & Katzenstein, 2006; Beutler & Harwood, 2000; Connolly Gibbons, 2003; Kendall, Chu, Gifford, Hayes, & Nauta, 1998; Luborsky, 1993), the NT therapeutic manual was developed to be sensitive to each client's progress during treatment, adjusting specific interventions to the idiosyncratic characteristics of the clients. For the strategies and techniques, we refer the reader to the original unpublished manual (Gonçalves & Bento, 2008, upon request) and to previous work by White (2007; White & Epston, 1990), upon which the manual was based.

Cognitive-Behavioral Therapy. CBT treatment relies on the principle that depressive symptoms are maintained due to a dysfunctional way of interpreting reality. The client is therefore encouraged to attempt new ways of thinking about himself or herself, the world, and/or others and to test these newly reformulated hypotheses in reality. Because

the literature about CBT for depression is vast, we refer the reader to the original treatment manual upon which our treatment was based (Beck et al., 1979; Leahy & Holland, 2000).

Adherence to the manual and competence. Therapists chose the model they wanted to use given their prior training or preferences. Treatment integrity was ensured through supervision and the assessment of adherence to the therapeutic manuals from the perspective of external judges. At the beginning of the study, the two graduate psychotherapists received weekly supervision with two seasoned psychotherapists who had extensive experience as supervisors. Later in the study, they helped the graduate psychotherapists (the first and fourth authors) supervise the younger therapists who joined the project. All sessions were videotaped and watched during the supervision meetings so that the supervisors could keep track of adherence to the manual and the quality of the interventions. After the study was completed, videotapes of sessions 4, 8, and 12 of 25% of the randomly selected cases were rated for adherence to the manual by external raters. Each rater rated one case using the ACS-N-CBT scale (Gonçalves, Bento, et al., 2009). The raters ($n = 10$) were all psychologists and experienced psychotherapists who were familiar with both NT and CBT theory and practice (mean age = 32.50, SD = 8.38; mean years of clinical experience = 7.80, SD = 5.49) and were not involved in the trial.

Analyses

Following the Consolidated Standards of Reporting Trials (Moher, Schulz, Altman, & CONSORT Group, 2001), separate intent-to-treat and completer analyses were performed.

Intent-to-treat group and missing data. The intent-to-treat (ITT) analysis included all clients who began treatment. To address the missing data in the case of the clients who left treatment prematurely, the method of carrying the last observation forward (LOCF) was used.

Statistical analysis. *Group differences*. The numbers of clients who dropped out of each treatment group were compared using a chi-square test. Other differences between groups on categorical variables were evaluated using chi-square tests. Independent-samples t-tests were used to test the initial differences between the two groups on continuous variables.

Efficacy. Cohen's d was computed to express the effect size of the pre- to post-treatment change in each treatment condition. An analysis of covariance (ANCOVA) was conducted for each of the outcome measures to test the differences between treatments in reducing depressive symptomatology, with the last observation as the dependent variable, the treatment group as the independent variables, and the first observation of the outcome variable as the covariate.

Clinical significance. To examine the variety of individual responses to treatment conditions, clinical significance was assessed according to Jacobson and Truax's (1991) criteria: (a) the change should be reliable (that is, greater than the RCI for the respective measure) and (b) at the end of the treatment, clients should move from a dysfunctional population range to a range typical of the functional population. We also present the proportion of clients who responded according to the Percentage Improvement method (PI). This method defines treatment response as a pre- to post-treatment change greater than 50% of the client's dysfunctional score. Following recommendations to use the cut-off scores generated from a meta-analysis (Hiller, Schindler, & Lambert, 2012), for the BDI-II, we used a normative value of 14.29 points to separate the non-clinical and clinical population, as suggested by Seggar et al. (2002). For example, if a client scored 30 at pre-treatment, the number of points to be considered 100% of the change was above 14.29, or 15.71 points. Because 50% of 15.71 is 7.88, any change greater than this value was considered a response for this particular client.

Rate of change. A Hierarchical Linear Model (HLM) analysis was used to assess whether the two treatments, NT and CBT, differed in the pace of symptom reduction over time. The analysis was conducted separately for the BDI-II and the OQ-45.2 scores. All measurements for each subject were used regardless of whether the participants completed treatment or were considered dropouts. At level 1, the within-subject level, we modeled the individual slope and intercept of the outcome measures for each participant over time (reflecting each individual's starting point and rate of change on either the BDI-II or the OQ-45.2). At level 2, the between-subject level, the aforementioned participant-specific parameters (slope and intercept) were modeled using only the group assignment to test whether differences in the rate of change could be predicted by it.

Results

Preliminary Analyses

Participants' flow. During the two and a half years it took to complete the trial (from February 2008 to October 2010), 107 people were assessed for eligibility, 26 of whom did not meet the inclusion criteria (see Figure 1). Of the 81 subjects who were assigned to treatments, 16 refused to participate in the study or did not return after the initial assessment, and two were excluded for Axis-II co-morbidity detected during treatment. These clients were referred to more adequate treatment in the clinic. Of the remaining 63 clients, 34 were allocated to NT and 29 to CBT. Of the 34 participants who received NT, 14 dropped out. Of the 29 participants who received CBT, nine dropped out.

Comparison between participants at baseline. Baseline demographic variables and clinical characteristics for the participants of each treatment condition are presented in Table I according to the CONSORT statement guidelines (Moher et al., 2001). No meaningful differences were found between the participants in the two treatment conditions. As shown in Table I, the participants did not differ significantly on any of the demographic variables at the time of intake except for the number of years of education ($NT_{Mean} = 14.41$, $SD = 4.07$; $CBT_{Mean} = 12.03$, $SD = 5.11$; $t(61) = 2.05$, $p < .044$). There was also no major group difference (Table I) in any of the clinical variables (global functioning, presence of co-morbidity, intake of

medication, previous hospitalizations, previous suicide attempts, and previous experiences with psychotherapy). GAF (DSM-IV; American Psychiatric Association, 2000) was approximately 60 ($NT_{Mean} = 58.94$, $SD = 10.87$; $CBT_{Mean} = 60.90$, $SD = 10.06$; $t(61) = -.73$, $p = .465$), which corresponds to moderate symptoms and moderate impact on social or occupational functioning. Still regarding initial severity, neither on the BDI-II ($NT_{Mean} = 29.08$, $SD = 9.62$; $CBT_{Mean} = 33.89$, $SD = 11.46$; $t(61) = -1.80$, $p = .075$) nor on the OQ-45.2 ($NT_{Mean} = 92.14$, $SD = 15.55$; $CBT_{Mean} = 97.03$, $SD = 22.51$; $t(61) = -.73$, $p = .31$) were differences found. Most clients had no co-morbid anxiety disorder ($NT = 85.3\%$; $CBT = 72.4\%$, $\chi^2(1) = 1.58$, $p = .20$) and had never been hospitalized ($NT = 94.1\%$; $CBT = 82.8\%$, $\chi^2(1) = 2.04$, $p = .15$). In both treatments, a considerable number of clients were taking medication at the beginning of treatment ($NT = 64.7\%$; $CBT = 55.2\%$, $\chi^2(1) = 0.59$, $p = .44$). Several clients had previous suicide attempts ($NT = 11.8\%$; $CBT = 20.7\%$, $\chi^2(1) = 0.93$, $p = .33$), and a few had previous experiences with psychotherapy ($NT = 11.8\%$; $CBT = 13.8\%$, $\chi^2(1) = 0.06$, $p = .81$).

Number of sessions in each treatment. Of the 40 completers, 24 clients received 20 sessions of treatment, and 16 clients completed treatment between the eighth and 19th sessions. NT clients, including dropouts, received an average of 12.94 sessions ($SD = 7.05$), whereas CBT clients received an average of 14.90 sessions ($SD = 6.48$). This difference was not statistically significant
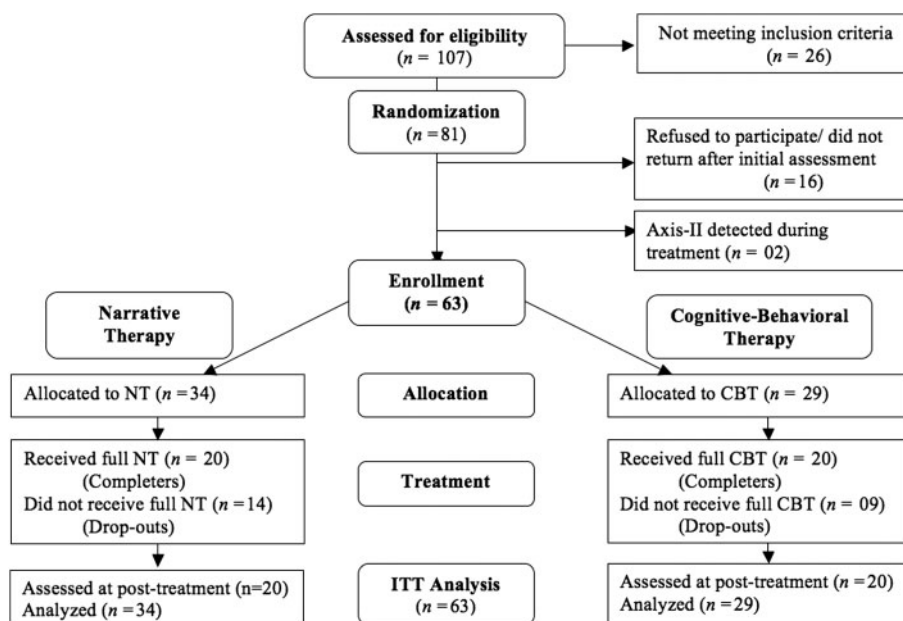

Figure 1. CONSORT flow chart in the clinical trial.
Note. NT = Narrative Therapy; CBT = Cognitive-Behavioral Therapy; ITT = Intend-to-treat.

Table I. Clinical and demographic characteristics of the sample at baseline, according to treatment condition.

Variable		NT (n = 34)		CBT (n = 29)		t (61)	χ^2 (1)
		M (SD)	n (%)	M (SD)	n (%)		
Age		37.18 (12.72)		33.41 (9.72)		1.30	
Gender	Female		27 (79)		24 (82.8)		0.11
	Male		7 (21)		5 (17.2)		
Education (in years)		14.41 (4.07)		12.03 (5.11)		−2.05★	
SES (Graffar[1])		12.25 (2.75)		12.61 (3.75)		0.54	
Relationship	Single		15 (44.1)		12 (41.4)		
Status	Married		10 (29.4)		11 (37.9)		2.07
	Divorced		7 (20.6)		6 (20.7)		
	Widowed		2 (5.9)		0 (0)		
Professional	Employed		14 (41.2)		18 (62.1)		
Status	Unemployed		10 (29.4)		4 (13.8)		4.8
	Student		8 (23.5)		7 (24.1)		
	Retired		2 (5.9)		0 (0)		
GAF at intake		58.94 (10.87)		60.90 (10.06)		−.73	
Score BDI-II Session 01		29.08 (9.62)		33.89 (11.4)		−1.80†	
Score OQ.45.2 Session 01		92.14 (15.5)		97.03 (22.5)		−1.01	
Co-morbidity at intake	No		29 (85.3)		21 (72.4)		
(anxiety problems)	Yes		5 (14.7)		8 (27.6)		1.58
Psychiatric medication at intake	No		12 (35.3)		13 (44.8)		
	Yes		22 (64.7)		16 (55.2)		0.59
Previous hospitalizations	No		32 (94.1)		24 (82.8)		
	Yes		2 (5.9)		5 (17.2)		2.04
Previous suicide attempts	No		30 (88.2)		23 (79.3)		
	Yes		4 (11.8)		6 (20.7)		0.93
Previous experiences with	No		30 (88.2)		25 (86.2)		
psychotherapy	Yes		4 (11.8)		4 (13.8)		0.06

[1] Graffar is an international classification for socioeconomic status (Graffar, 1956). The higher the score is, the lower the SES. This sample means (NT M = 12.2; CBT M = 12.6) indicate high SES.
Note. SES = socioeconomic status; GAF = Global Assessment of Functioning of the DSM-IV (First et al., 2002); BDI-II = Beck Depression Inventory II; OQ-45.2 = Outcome Questionnaire 45.2
★ Indicates statistical significance at $p < .05$; † indicates marginally significant difference at $p < .10$.

(t (61) = −1.13, p = .278). Termination occurred either at the end of the treatment manual (20 sessions, n = 24) or at a time that was mutually agreed between the therapist and the client at an earlier stage of the treatment manual (ranging from eight to 19 sessions, n = 16).

Adherence and competence assessment. External raters (n = 16) watched videotapes of sessions 4, 8 and 12 of 25% of the cases (n = 16) and rated the sessions using the ACS-N-CBT (Gonçalves, Bento, et al., 2009). As expected, in the NT treatment condition, therapists used significantly more NT techniques than in the CBT treatment condition (U = 0.500, p < .001). Comparably, in the CBT treatment condition, CBT therapists used significantly more CBT techniques than in the NT treatment condition (U = 0.00, p < .001). The use of general therapeutic attitudes did not differ across treatments (U = 264, p = .307). General and specific competence rated by external observers did not differ between the two treatments (U = 257.5, p = .692 and U = 206, p = .136, respectively).

Influence of medication on outcome. To determine whether the improvement of clients taking medication differed on the BDI-II compared to clients not taking medication, we used an ANCOVA. The initial BDI-II score was used as the covariate, the final BDI-II score was the dependent variable, and medication at intake was an independent variable. There was no significant effect of medication at intake on the BDI-II ($F(1,60)$ = 0.14, p = .71). A similar analysis conducted with the OQ-45.2 also revealed no significant effect for medication at intake ($F(1,60)$ = 1.21, p = .28).

Influence of therapist's experience on outcome. Because the number of clients seen by novice therapists in CBT was significantly higher than the number of clients seen by novice therapists in NT, we tested the effects of the therapist's experience (novice, apprentice, or graduate) on the final score on the BDI-II using an ANCOVA, with the initial BDI-II score as the covariate. No significant effect was found for therapist experience on the outcome ($F(2,59)$ = 0.78, p = .463).

Symptom reduction during treatment

Pre- to post-treatment means, standard deviations, and effect sizes for both treatment groups are presented in Table II. CBT had higher pre- to post-treatment effect sizes when compared to NT on the BDI-II, both for ITT ($d = 1.25$ for CBT and $d = 0.56$ for NT) and completer ($d = 1.51$ for CBT and $d = 0.89$ for NT) analysis. For OQ-45.2, CBT and NT pre- to post-treatment effect sizes were very similar (ITT analysis: $d = 0.65$ for CBT and $d = 0.69$ for NT; completer analysis: $d = 0.91$ for CBT and $d = 0.84$ for NT).

ANCOVAs comparing the observed post-treatment change with the pre-treatment scores as covariates suggest that both treatment groups showed significant improvement in all measures for both the ITT ($n = 63$) and the completer ($n = 40$) analyses. Comparing CBT and NT, a significant group effect was found for the BDI-II ($F(1,60) = 4.16$, $p = .046$) and its somatic dimension ($F(1,60) = 4.04$, $p = .049$), but only for the entire ITT sample. No significant group effects for the BDI-II were found for the completer analysis. Additionally, no significant group differences were found for the OQ-45.2 for the ITT analysis ($F(1,60) = 0.15$, $p = .70$) or the completer analysis ($F(1,37) = 0.04$, $p = .83$). Moreover, separate ANCOVAs were conducted for each of the three subscales of the OQ-45.2, and no difference in the amount of change was found between any groups (for Symptom Distress, $F(1,60) = 0.017$, $p = .90$; for Interpersonal Relatedness, $F(1,60) = 0.78$, $p = .38$; and for Social Role Functioning, $F(1,60) = 0.073$, $p = .79$).

Clinical significance. Each client's post-treatment score was examined to determine (a) whether it reliably changed, (b) whether it fell below the cutoff scores for the functional distributions (Functional Population, also referred to as remission; Hiller et al., 2012), (c) whether it moved to the functional distribution and improved reliably at the same time (Clinical Significant Change), (d) whether the final score increased more than the RCI for the specific measure (i.e., whether individuals reliably deteriorated during treatment), and (e) whether it decreased by more than 50% of its value on the dysfunctional range of the scale (i.e., whether clients responded). Table III shows the results for the ITT and completer analysis separately.

In the completer analysis, the group proportions on the BDI-II were very similar on all of the categories of clinical change (i.e., the proportion of clients who attained reliable improvement (CBT = 75%, NT = 50%; $\chi^2 (1) = 2.667$, $p = .102$), who moved to the functional population (CBT = 50%, NT = 45%; $\chi^2 (1) = 0.1$, $p = .751$), or who attained clinically significant change (CBT = 50%, NT = 40%; $\chi^2 (1) = 0.404$, $p = .525$)). However, in the ITT analysis, the proportion of clients that improved reliably was marginally larger for the CBT group (CBT = 62.1%, NT = 38.2%; $\chi^2 (1) = 3.55$, $p = .059$). The proportion of responders (according to the PI method) was significantly larger in the CBT group (CBT = 68.9%, NT = 41.2%; $\chi^2 (1) = 4.865$, $p = .027$).

Rate of change across treatment conditions. The results from the HLM comparing the rates of change in the two groups indicate that according to the BDI-II, group allocation was not significantly related to individual growth rates (coefficient of BDI-II for group: $b = 0.283$, $t(62) = 1.442$, $p = 0.154$). This was also the case using the OQ-45.2 (coefficient of OQ-45.2 for group: $b = 0.163$,

Table II. Intent-to-treat and completer analysis pre- and post-treatment mean scores, standard deviations, and effect sizes for NT and CBT.

| | Treatment condition | | | | | |
| | NT ($n = 34$) | | | CBT ($n = 29$) | | |
Measure	Pre Mean (SD)	Post Mean (SD)	d^1 (CI)	Pre	Post	d (CI)
	Intent-to-treat (LOFC, $N = 63$)					
BDI-II	29.09 (10)	22.59 (13)	0.56 (0.08–1.05)	33.90 (11)	18.90 (13)	1.25 (0.68–1.81)
OQ-45.2	92 (16)	75.91 (29.1)	0.69 (.2–1.17)	97 (23)	79.10 (31.3)	0.65 (0.12–1.18)
	Completer ($n = 40$)					
	NT ($n = 20$)			CBT ($n = 20$)		
BDI-II	29.4 (10.4)	18.8 (13.2)	0.89 (0.24–1.54)	33.5 (9.9)	16.4 (12.6)	1.51 (0.81–2.21)
OQ-45.2	(4) 94 (17)	(5) 71.8 (33.1)	0.84 (0.20–1.49)	97.7 (23.2)	72.2 (32.4)	0.91 (0.26–1.56)

Note. NT = Narrative Therapy; CBT = Cognitive-Behavioral Therapy; BDI-II = Beck Depression Inventory-II; OQ-45.2 = Outcome Questionnaire 45.2; SD = standard deviation; LOFC = last observation carried forward; CI = Confidence interval.
[1] Effect sizes refer to pre- and post-treatment and are categorized along a continuum of "no effect" (ES < 0.2), "small effect" (0.2 ≤ ES ≤ 0.5), "medium effect" (0.5 ≤ ES ≤ 0.8), and "large effect" (ES ≥ 0.8) (Cohen, 1988).

Table III. The percentage of reliable improvement, movement to functional population, clinically significant change, and reliable deterioration by treatment condition and outcome measure.

Type of change	BDI-II		OQ-45.2	
	NT n (%)	CBT n (%)	NT n (%)	CBT n (%)
ITT (n = 63)				
Reliable Improvement[a]	13 (38.2)	18 (62.1)[†]	12 (35.3)	14 (48.2)
Functional Population / Remission[b]	11 (32.4)	12 (41.4)	10 (29.4)	8 (27.6)
Clinical Significant Change[c]	10 (29.4)	12 (41.4)	9 (26.5)	6 (20.7)
Reliable Deterioration[d]	1 (2.9)	0 (0)	1 (2.9)	5 (17.24%)
Response[e]	14 (41.2)	20 (68.9)[*]	14 (41.2)	14 (48.3)
Completers (n = 40)				
Reliable Improvement	10 (50)	15 (75)	12 (35.3)	11 (38)
Functional Population	9 (45)	10 (50)	9 (45)	7 (35)
Clinical Significant Change	8 (40)	10 (50)	9 (45)	5 (25)
Reliable Deterioration	0 (0)	0 (0)	1 (2,9)	3 (10)
Response	11 (55)	15 (75%)	13 (65)	12 (60)

Note. BDI-II = Beck Depression Inventory-II; OQ-45.2 = Outcome Questionnaire-45.2; NT = Narrative Therapy; CBT = Cognitive Behavioral Therapy; ITT = Intend-to-treat; [†] = $p < .10$ and [*] = $p < .05$.
[a] Reliable improvement = proportion of clients who attained Reliable Improvement (changed more than the Reliable Change Index: for BDI-II = 8.46; for OQ-45.2 = 15); [b] Functional Population / Remission = proportion of clients who moved into the Functional Population (below the 80th percentile; for BDI-II = 14.29; for OQ-45.2 = 62); [c] Clinical Significant Change = clients who simultaneously showed reliable improvement AND moved into the Functional Population; [d] proportion of clients who deteriorated reliably (more than Reliable Change Index); [e] Response = change greater than 50% of pre-treatment score, considering the number of points situated on the dysfunctional range of the scale (Percent Improvement Method; Hiller et al., 2012).

$t(62) = 0.33$, $p = 0.742$). The rate of symptom reduction for the BDI-II was 0.84 points per session for the CBT group and 0.56 points for the NT group. For the OQ-45.2, the rate of symptom reduction was 1.17 points per session for the CBT group and 1.01 points for the NT group. Figure 2 illustrates the rate of change for both groups.

Dropout across treatment conditions. Dropout in our study was defined as unilateral termination by the client without the therapist's approval or knowledge (Richmond, 1992) and/or failure to attend the last scheduled appointment (Hatchett, Han, & Cooker, 2002; Hatchett & Park, 2003). Fourteen clients (41.2%) dropped out in the NT group compared to nine (31%) in the CBT group, but this difference was not statistically significant (Fisher's exact probability = .27). The overall dropout rate was 36.50%.

NT and CBT dropouts did not differ on any of the baseline demographic variables (age, gender, relationship status, professional status, years of education, and socioeconomic status) or on clinical baseline variables (severity of depression, presence of co-morbidity, medication at intake, previous hospitalizations, and previous experiences with psychotherapy).
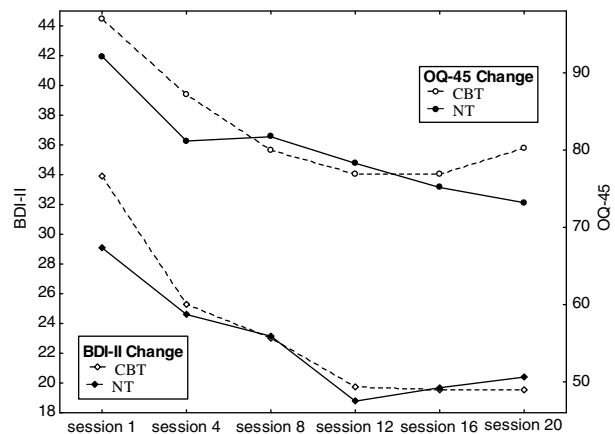

Figure 2. Course of symptom reduction in both treatment conditions on BDI-II and OQ-45.2.
Note. BDI-II = Beck Depression Inventory-II; OQ-45.2 = Outcome Questionnaire-45.2, NT = Narrative Therapy; CBT = Cognitive Behavioral Therapy.

Discussion

The present study aimed to assess the efficacy of NT for major depression in adults, a popular treatment among clinicians that has very little empirical evidence in its favor, compared with CBT. The results showed that by the end of treatment, clients in both treatment groups scored significantly lower on all measures. Although the completer analyses yielded no evidence indicating differential effects or the superiority of either CBT or NT in reducing depressive symptoms (BDI-II) or general psychosocial symptoms and problems (OQ-45.2), surprisingly, when the analysis included dropouts (i.e., ITT analysis), the CBT group showed a significantly larger decrease of depressive symptoms at post-treatment and a significantly larger proportion of

treatment responders. This difference may indicate that dropouts in the CBT group have better gains in reducing depressive symptoms than the NT dropouts. One possible clinical explanation for this difference is the inclusion in the CBT manual of a clear and structured behavioral activation module at the beginning of treatment, in which therapists have the opportunity to work exclusively on the alleviation of depressive symptoms. In contrast, in NT, therapists directly begin to intervene with a focus on the narrative change. NT intervention may require more cognitive resources than some depressive clients have available at the onset of treatment. If this hypothesis is confirmed in future studies of NT, narrative therapists should consider ways to address this problem, including a more supportive module at the beginning of the treatment manual or explaining that the client might feel that his condition is worsening before he begins to get better, as is often the case with antidepressant medication. Warning the client about this effect may change the client's expectation of the course of treatment and thus enhance adherence to it.

The better results on the BDI-II for the ITT sample may not be surprising given that CBT treatment specifically aims to reduce depressive symptoms in the same way that the BDI-II aims to measure depressive symptoms. Future research should consider the use of measures focused on narrative change (e.g., innovative moments; Gonçalves, Matos, & Santos, 2009; Gonçalves, Ribeiro, Mendes, Matos, & Santos, 2011) and outcome measures using independent raters (e.g., the Hamilton-D scale; Fleck, Poirier-Littre, Guelfi, Bourdel, & Loo, 1995) to assess whether NT produces changes that are more consistent with its theoretical background.

An established benchmark for the psychotherapeutic treatment of depression outcomes (Minami, Wampold, Serlin, Kircher, & Brown, 2007) provided parameters for the most efficacious treatments (efficacy benchmark) and parameters for clients who were randomly assigned to the waiting list group control (natural history benchmark). Comparing the natural history benchmark with the pre-post effect size yielded in the ITT analyses ($d = 0.56$) we conclude that there is insufficient evidence to support NT as an efficacious treatment for depression. In contrast, the effect size in the completer analysis of NT ($d = 0.89$) is significantly greater than the natural history. Thus, completing NT treatment is reliably better than no treatment. However, even for the completers, the effect size is not large enough to conclude that the treatment is clinically equivalent to the most efficacious treatments in clinical trials. Caution must be taken when comparing our results

with this benchmark given that our sample presented higher initial severity at the BDI-II than the sample used to calculate this benchmark (Minami et al., 2007). Benchmarking CBT data against the values provided by Minami et al. (2007) shows that undergoing CBT is clinically superior than no treatment for the more conservative ITT analysis ($d = 1.25$) and the less conservative completer analysis ($d = 1.51$). Whereas for the ITT analysis, the effect size is not large enough to state that the CBT group was as effective as the efficacy benchmark, for the completer analysis it is very likely that the CBT provided in this trial was as effective as any effective treatment.

As mentioned previously, Vromans and Schweitzer (2011) have recently conducted a clinical trial with NT, the only empirical study that is directly comparable to ours. For those who complete treatment, BDI-II pre-post effect sizes are comparable to our findings. However, the ITT pre-post effect size comparison confirms that our sample had poorer improvement. Future studies may clarify this issue, especially given that Vromans and Schweitzer (2011) used a much shorter intervention (eight sessions, compared to 20 sessions in the current study).

The overall dropout rate was 35.5%. Considering the estimate of the weighted mean dropout rate of 19.7% obtained by a meta-analysis including a large number of studies (Swift & Greenberg, 2012), this is a high dropout rate. The considerable number of young therapists in our sample might explain this rate. There is some evidence suggesting a relationship between the level of the therapist's experience and dropout rates. Trainees who had not yet obtained their degrees have higher dropout rates (27%) than trainees who had already obtained their degrees (17%; Swift & Greenberg, 2012). Similarly, some evidence has shown significantly higher dropout rates in university clinics (Kadera, Lambert, & Andrews, 1996; Swift, Callahan, Heath, Herbert, & Levine, 2010).

One might argue that the fact that our sample included clients on medication may create a bias favoring the efficacy of the treatments. However, the outcomes of clients taking medication and clients not taking medication did not differ. In addition, the proportion of medicated clients was evenly distributed across treatments, which should attenuate any possible bias.

Strengths and Limitations

The strengths of the study included the following: (1) the sample size, which satisfied the analysis we applied and is comparable to most studies in the area (Chambless & Hollon, 1998; Kazdin & Bass, 1989); (2) the assessment of clinically significant change; (3)

the presentation of the ITT and completer analyses separately adds information on the dropouts during treatment; (4) the use of manualized treatments favors the principle of replicability (Chambless & Hollon, 1998); and (5) the integrity assessment indicates that therapists followed the manual and that they did so with some degree of competence, which increases confidence in the results.

Following the criteria to identify ESTs (Chambless & Hollon, 1998), our study used a controlled clinical trial design to test the efficacy of NT, the therapists followed treatment manuals, and the characteristics of the sample were clearly specified and thoroughly assessed. Although these characteristics can be regarded as strengths, the question may be raised of the generalizability of the results. The sample included in this trial was very homogeneous and may not be the population normally found in clinical settings (i.e., clients with multiple co-morbid Axis I and II conditions). This methodological choice emphasizes the internal validity and sacrifices the external validity of the results (Hansen, Lambert, & Forman, 2002; Lambert & Ogles, 2004).

The current study has other limitations. The diversity in the therapists' level of expertise may have been a confounding variable because many cases were seen by novice therapists and were not evenly distributed in both conditions (the CBT group had significantly more clients seen by novice therapists than the NT group). However, the impact of this limitation on the results is unclear because the level of therapist experience was not related to the outcome, a finding similar to that of other researchers (e.g., Kadera et al., 1996). The fact that the therapists did not see pilot cases before seeing the actual clients for the trial may also be considered a limitation. The nested design used in this study (therapists were divided by treatment group) increases the possibility of therapist effects.

Another limitation was the way the clients were assigned to treatments. Every other patient was assigned to either CBT or NT as they arrived in the clinic. It is not clear how this method may interfere with the results because (a) this method was conducted in a systematic manner by a clinic administrative staff member who was unaware of the research design, and (b) there was most likely not a specific order or pattern in which clients sought treatment. Although this may technically be considered a quasi-random method, it is still worth noting the possibility of bias.

Implications of the Study

As discussed, the data were not consistent enough to ascertain the efficacy of Narrative Therapy for moderate depression, especially due to a high number of client dropouts who did not seem to benefit from treatment. However, when clients completed treatment, it appears to be an effective treatment. Thus, the main conclusion of this study is that additional efficacy studies are needed before NT can be recommended for the treatment of depression.

References

Ablon, J. S., Levy, R. A., & Katzenstein, T. (2006). Beyond brand names of psychotherapy: Identifying empirically supported change processes. *Psychotherapy: Theory, Research, Practice, Training*, 43, 216–231. doi:10.1037/0033-3204.43.2.216

American Psychiatric Association. (2000). *DSM-IV-TR*: *Diagnostic and Statistical Manual of Mental Disorders* (4th rev. ed.). Washington, DC: American Psychiatric Press.

Andrews, P. W., & Thomson, J. A. (2009). The bright side of being blue: Depression as an adaptation for analyzing complex problems. *Psychological Review*, 116, 620–654. doi:10.1037/a0016242

Angus, L. E., & McLeod, J. (2004). *The handbook of narrative and psychotherapy*: *Practice, theory and research*. Los Angeles, CA: Sage.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II (BDI-II)*. San Antonio, TX: Psychological Corporation.

Betchley, D., & Falconer, W. (2002). Giving Colin a voice: The challenge of narrative therapy with a client who has an intellectual disability and communication difficulties. *Australian Social Work*, 55, 3–12. doi:10.1080/03124070208411667

Beutler, L. E., & Harwood, T. M. (2000). *Prescriptive psychotherapy: A practical guide to systematic treatment selection*. Oxford: Oxford University Press.

Busch, R. S. (2011). *Contextualising a problematic relationship between narrative therapy and evidence-based psychotherapy evaluation in psychology*. Massey University, Palmerston North, New Zealand.

Butler, A. C., Chapman, J. E., Forman, E. M., & Beck, A. T. (2006). The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clinical Psychology Review*, *26*, 17–31. doi:10.1016/j.cpr.2005.07.003

Campos, R. C., & Gonçalves, B. (2011). The Portuguese version of the Beck Depression Inventory-II (BDI-II). *European Journal of Psychological Assessment*, *27*, 258–264. doi:10.1027/1015-5759/a000072

Cashin, A. (2008). Narrative Therapy: A psychotherapeutic approach in the treatment of adolescents with Asperger's disorder. *Journal of Child and Adolescent Psychiatric Nursing*, *21*, 48–56. doi:10.1111/j.1744-6171.2008.00128.x

Chambless, D. L., Baker, M. J., Baucom, D. H., Beutler, L. E., Calhoun, K. S., Crits-Christoph, P., … Woody, S. (1998). Update on empirically validated therapies, II. *Clinical Psychologist*, *51*, 3–16.

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, *66*, 7–18. doi:10.1037/0022-006X.66.1.7

Chenail, R. J., DeVicentis, M., Kiviat, H., & Somers, C. (2012). A systematic narrative review of discursive therapies research: Considering the value of circumstantial evidence. In A. Lock & T. Strong (Eds.), *Discursive perspectives in therapeutic practice*. Oxford: Oxford University Press.

Coelho, R., Martins, A., & Barros, H. (2002). Clinical profiles relating gender and depressive symptoms among adolescents ascertained by the Beck Depression Inventory II. *European Psychiatry*, *17*, 222–226. doi:10.1016/S0924-9338(02)00663-6

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Connolly Gibbons, M. B. (2003). Flexibility in manual-based psychotherapies: Predictors of therapist interventions in interpersonal and cognitive-behavioral therapy. *Psychotherapy Research*, *13*, 169–185. doi:10.1093/ptr/kpg017

Cuijpers, P., van Straten, A., Andersson, G., & van Oppen, P. (2008). Psychotherapy for depression in adults: A meta-analysis of comparative outcome studies. *Journal of Consulting and Clinical Psychology*, *76*, 909–922. doi:10.1037/a0013075

Da Costa, D., Nelson, T. M., Rudes, J., & Guterman, J. G. (2007). A narrative approach to body dysmorphic disorder. *Journal of Mental Health Counseling*, *29*, 67–80.

De Jong, K., Nugter, M. A., Polak, M. G., Wagenborg, J. E. A., Spinhoven, P., & Heiser, W. J. (2007). The outcome questionnaire (OQ-45) in a Dutch population: A cross-cultural validation. *Clinical Psychology & Psychotherapy*, *14*, 288–301. doi:10.1002/cpp.529

DeRubeis, R. J., & Crits-Christoph, P. (1998). Empirically supported individual and group psychological treatments for adult mental disorders. *Journal of Consulting and Clinical Psychology*, *66*, 37–52. doi:10.1037/0022-006X.66.1.37

Dobson, K. S. (1989). A meta-analysis of the efficacy of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology*, *57*, 414–419. doi:10.1037/0022-006X.57.3.414

Draucker, C. B. (1998). Narrative therapy for women who have lived with violence. *Archives of Psychiatric Nursing*, *12*, 162–168. doi:10.1016/S0883-9417(98)80018-6

Etchison, M., & Kleist, D. M. (2000). Review of Narrative Therapy: Research and utility. *The Family Journal*, *8*, 61–66. doi:10.1177/1066480700081009

First, M. B., Gibbon, M., Spitzer, R. L., Williams, J., & Benjamin, L. (1997). *Structured Clinical Interview for DSM-IV Axis II personality disorders: SCID-II*. Washington, DC: American Psychiatric Press.

First, M. B., Spitzer, R., Gibbons, M., & Williams, J. (2002). *Structured clinical interview for DSM-IV-TR Axis I Disorders, research version, non-patient edition (SCID-I/NP)*. New York: Biometrics Research, New York State Psychiatric Institute.

Fleck, M. P. A., Poirier-Littre, M. F., Guelfi, J.-D., Bourdel, M. C., & Loo, H. (1995). Factorial structure of the 17-item Hamilton Depression Rating Scale. *Acta Psychiatrica Scandinavica*, *92*, 168–172. doi:10.1111/j.1600-0447.1995.tb09562.x

Freedman, J., & Combs, G. (1996). *Narrative therapy: The social construction of preferred realities*. New York: WW Norton.

Gloaguen, V., Cottraux, J., Cucherat, M., & Blackburn, I. M. (1998). A meta-analysis of the effects of cognitive therapy in depressed patients. *Journal of Affective Disorders*, *49*, 59–72. doi:10.1016/S0165-0327(97)00199-7

Gonçalves, M. M., & Bento, T. (2008). *Manual terapêutico psicoterapia narrativa de re-autoria*. Braga, Portugal.

Gonçalves, M. M., Bento, T., Lopes, R., & Salgado, J. (2009). *Manual da escala de adesão e competência*. Braga, Portugal.

Gonçalves, M. M., Matos, M., & Santos, A. (2009). Narrative therapy and the nature of "innovative moments" in the construction of change. *Journal of Constructivist Psychology*, *22*, 1–23. doi:10.1080/10720530802500748

Gonçalves, M. M., Ribeiro, A. P., Mendes, I., Matos, M., & Santos, A. (2011). Tracking novelties in psychotherapy process research: The innovative moments coding system. *Psychotherapy Research*, *21*, 497–509. doi:10.1080/10503307.2011.560207

Graffar, M. (1956). Une méthode de classification sociale d'echantillons de population. *Courrier*, *6*, 455–59.

Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose–response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice*, *9*, 329–343. doi:10.1093/clipsy.9.3.329

Hatchett, G. T., Han, K., & Cooker, P. G. (2002). Predicting premature termination from counseling using the Butcher treatment planning inventory. *Assessment*, *9*, 156–163. doi:10.1177/10791102009002006

Hatchett, G. T., & Park, H. L. (2003). Comparison of four operational definitions of premature termination. *Psychotherapy: Theory, Research, Practice, Training*, *40*, 226–231. doi:10.1037/0033-3204.40.3.226

Hayes, A., & Strunk, D. (n.d.). Research-supported psychological treatments for depression. *Society for Clinical Psychology - American Psychological Society, Division 12*. Retrieved December 5, 2012, from http://www.div12.org/PsychologicalTreatments/disorders/depression_main.php.

Hiller, W., Schindler, A. C., & Lambert, M. J. (2012). Defining response and remission in psychotherapy research: A comparison of the RCI and the method of percent improvement. *Psychotherapy Research*, *22*, 1–11. doi:10.1080/10503307.2011.616237

Jacobson, N. S., & Truax, P. (1991). Clinical Significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*, 12–19. doi:10.1037/0022-006X.59.1.12

Kadera, S. W., Lambert, M. J., & Andrews, A. A. (1996). How much therapy is really enough? A session-by-session analysis of the psychotherapy dose-effect relationship. *Journal of Psychotherapy Practice & Research*, *5*, 132–151.

Kazdin, A. E., & Bass, D. (1989). Power to detect differences between alternative treatments in comparative psychotherapy outcome research. *Journal of Consulting and Clinical Psychology*, *57*, 138–147. doi:10.1037/0022-006X.57.1.138

Kendall, P. C. (1998). Empirically supported psychological therapies. *Journal of Consulting and Clinical Psychology*, *66*, 3–6. doi:10.1037/0022-006X.66.1.3

Kendall, P. C., Chu, B., Gifford, A., Hayes, C., & Nauta, M. (1998). Breathing life into a manual: Flexibility and creativity with manual-based treatments. *Cognitive and Behavioral Practice*, *5*, 177–198. doi:10.1016/S1077-7229(98)80004-7

Kendall, P. C., Holmbeck, G., & Verduin, T. (2004). Methodology, design, and evaluation in psychotherapy research. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 16–43). New York: Wiley.

Kropf, N., & Tandy, C. (1998). Narrative therapy with older clients: The use of a "meaning-making" approach. *Clinical Gerontologist*, *18*, 3–16. doi:10.1300/J018v18n04_02

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*, 249–258. doi:10.1002/(SICI)1099-0879(199612)3:4%3C249::AID-CPP106%3E3.0.CO;2-S

Lambert, M. J., & Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 139–193). New York: Wiley.

Leahy, R. L., & Holland, S. J. (2000). *Treatment plans and interventions for depression and anxiety disorders*. New York: Guilford Press.

Luborsky, L. (1993). Recommendations for training therapists based on manuals for psychotherapy research. *Psychotherapy: Theory, Research, Practice, Training*, *30*, 578–580. doi:10.1037/0033-3204.30.4.578

Machado, P. P., & Fassnacht, D. (2014) (Paper submitted for publication). *Normative data, reliability and clinical significance cut-offs for the Portuguese version of the Outcome Questionnaire (OQ-45)*.

Matos, M., Santos, A., Gonçalves, M. M., & Martins, C. (2009). Innovative moments and change in narrative therapy. *Psychotherapy Research: Journal of the Society for Psychotherapy Research*, *19*, 68–80. doi:10.1080/10503300802430657

Minami, T., Wampold, B. E., Serlin, R. C., Kircher, J. C., & Brown, G. S. (2007). Benchmarks for psychotherapy efficacy in adult major depression. *Journal of Consulting and Clinical Psychology*, *75*, 232–243. doi:10.1037/0022-006X.75.2.232

Moher, D., Schulz, K. F., Altman, D., & CONSORT Group. (2001). The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *Journal of the American Medical Association*, *285*, 1987–1991. doi:10.1001/jama.285.15.1987

Moreira, P., Beutler, L. E., & Gonçalves, Ó. (2008). Narrative change in psychotherapy: differences between good and bad outcome cases in cognitive, narrative, and prescriptive therapies. *Journal of Clinical Psychology*, *64*, 1181–1194. doi:10.1002/jclp.20517

Nylund, D. (2002). Poetic means to anti-anorexic ends. *Journal of Systemic Therapies*, *21*, 18–34. doi:10.1521/jsyt.21.4.18.23323

Orlinsky, D. E., & Rønnestad, M. H. (2005). *How psychotherapists develop: A study of therapeutic work and professional growth*. Washington, DC: American Psychological Association.

Palgi, Y., & Ben-Ezra, M. (2010). "Back to the future": Narrative treatment for post-traumatic, acute stress disorder in the case of paramedic Mr. G. *Pragmatic Case Studies in Psychotherapy*, *6*, 1–26.

Richmond, R. (1992). Discriminating variables among psychotherapy dropouts from a psychological training clinic. *Professional Psychology: Research and Practice*, *23*, 123–130.

Roness, A., Mykletun, A., & Dahl, A. A. (2005). Help-seeking behaviour in patients with anxiety disorder and depression. *Acta Psychiatrica Scandinavica*, *111*, 51–58. doi:10.1111/j.1600-0447.2004.00433.x

Rothschild, P., Brownlee, K., & Gallant, P. (2000). Narrative interventions for working with persons with AIDS. *Journal of Family Psychotherapy*, *11*, 1–13. doi:10.1300/J085v11n03_01

Rush, A. J., Beck, A. T., Kovacs, M., & Hollon, S. D. (1977). Comparative efficacy of cognitive therapy and pharmacotherapy in the treatment of depressed outpatients. *Cognitive Therapy and Research*, *1*, 17–37. doi:10.1007/BF01173502

Seggar, L. B., Lambert, M. J., & Hansen, N. B. (2002). Assessing clinical significance: Application to the Beck Depression Inventory. *Behavior Therapy*, *33*, 253–269. doi:10.1016/S0005-7894(02)80028-4

Steer, R. A., Brown, G. K., Beck, A. T., & Sanderson, W. C. (2001). Mean Beck Depression Inventory-II scores by severity of major depressive episode. *Psychological Reports*, *88*, 1075–1076.

Swift, J. K., Callahan, J. L., Heath, C. J., Herbert, G. L., & Levine, J. C. (2010). Applications of the psychotherapy phase model to clinically significant deterioration. *Psychotherapy: Theory, Research, Practice, Training*, *47*, 235–248. doi:10.1037/a0019787

Swift, J. K., & Greenberg, R. P. (2012). Premature discontinuation in adult psychotherapy: A meta-analysis. *Journal of Consulting and Clinical Psychology*. doi:10.1037/a0028226

Vromans, L. P., & Schweitzer, R. D. (2011). Narrative therapy for adults with major depressive disorder: Improved symptom and interpersonal outcomes. *Psychotherapy Research*, *21*, 4–15. doi:10.1080/10503301003591792

\White, M. (2007). *Maps of narrative practice*. New York: WW Norton.

White, M., & Epston, D. (1990). *Narrative means to therapeutic ends*. New York: WW Norton.

Young, E. (2008). Narrative therapy and elders with memory loss. *Clinical Social Work Journal*, *38*, 193–202. doi:10.1007/s10615-008-0146-4