IJTAR
International Journal of Transactional Analysis Research
8 (1), 3-20
https://doi.org/10.29044/v8i1p3

TA Treatment of Depression: A Hermeneutic Single-Case Efficacy Design Study - 'Anna'

© 2017 Enrico Benelli, Emanuela Moretti, Giorgio Cristiano Cavallero, Giovanni Greco, Vincenzo Calvo, Stefania Mannarini, Arianna Palmieri & Mark Widdowson

Abstract

This study is the first of a series of seven, and belongs to the second Italian systematic replication of findings from two previous series (Widdowson 2012a, 2012b, 2012c, 2013; Benelli, 2016a, 2016b, 2016c) that investigated the effectiveness of a manualised transactional analysis treatment for depression through Hermeneutic Single-Case Efficacy Design (HSCED). The therapist was a white Italian woman with 8 years of clinical experience and the client, Anna, was a 33-year old white Italian woman who attended 16 sessions of transactional analysis psychotherapy. Anna satisfied DSM-5 criteria for mild persistent depressive disorder (dysthymia) with anxious distress. The conclusion of the judges was that this was a good-outcome case: the dysthymic symptoms improved over the course of the therapy and were maintained in the 'healthy' range at the 6-month follow-up, the client reported a positive experience of the therapy and described important changes in intrapsychic and interpersonal patterns. In this case study, transactional analysis treatment for depression has proven its efficacy in treating persistent depressive disorder.

Key words

Systematic Case Study Research; Hermeneutic Single-Case Efficacy Design; Transactional Analysis Psychotherapy; Persistent Depressive Disorder (Dysthymia); Histrionic trait; Dependent trait.

Introduction

This study is the first of a series of seven, and belongs to the second Italian systematic replication of findings from two previous case series (Widdowson 2012a, 2012b, 2012c, 2013; Benelli, 2016a, 2016b, 2016c) and was conducted under the auspices of the European Association for Transactional Analysis (EATA) and the University of Padua.

Transactional analysis (TA) is a widely-practiced form of psyc are now what do hotherapy, supported with a vast literature (for a review see Ohlsson, 2010), but still it is under-recognised within the worldwide scientific community of psychotherapy. In order to define TA psychotherapy as an efficacious Empirically Supported Treatment (EST), its efficacy must have been established in at least one Randomised Clinical Trial (RCT) replicated by two independent research groups, or alternatively in at least three Single Case Experimental Design studies (SCED), replicated by at least two independent research groups, with each group conducting a case series of a minimum of three cases, without conflicting evidence (Chambless & Hollon, 1998). Recently, a wide community of researchers proposed that efficacy and effectiveness in psychotherapy are a complex object that cannot be adequately evaluated with either the experimental approach of RCT (Norcross, 2002; Westen, Novotny & Thomson-Brenner, 2004) or classical SCED (reverse or multiple baseline design) (McLeod, 2010). Systematic case study research has been proposed as a viable alternative to RCT and SCED (Iwakabe & Gazzola, 2009). Considering that approaches without evidence from RCTs tend to be under-recognised, Stiles, Hill and Elliott (2015) proposed collecting a series of mixed methods systematic single case studies as the first step toward recognition of marginalised and emerging models of psychotherapy.

Hermeneutic Single Case Efficacy Design (HSCED; Elliott, 2002; Elliott et al., 2009) is nowadays considered the most comprehensive set of methodological procedures for systematic case study research, and is a viable alternative to RCT and SCED in psychotherapy (McLeod, 2010). HSCED is gaining momentum with enhanced versions proposed by different research groups, to validate new psychotherapeutic approaches or extensions of previously validated psychotherapies for

investigation into their effectiveness with other disorders (e.g. Wall, Kwee, Hu & McDonald, 2016). Recently, a systematic review of all published HSCED studies found within English language peer-reviewed journals (Benelli, De Carlo, Biffi & McLeod, 2015) highlighted methodological issues related to different levels of stringency, offering solid alternatives to conducting sound research according to the available resources within practitioner research networks.

Systematic case study research has already been applied to investigate the effectiveness of TA for people with long term health conditions (McLeod, 2013a; 2013b) and HSCED methodology has been successfully applied to TA and widely described in this Journal by Widdowson (2012a). Recently, several HSCEDs supporting the effectiveness of TA treatment for depression (Widdowson, 2012a, 2012b, 2012c, 2013; Benelli, 2016a, 2016b, 2016c) have been published, as was an additional adjudicated study which demonstrated effectiveness of TA for mixed depression and anxiety (Widdowson, 2014). Furthermore, a related study was published on TA for emetophobia (Kerr, 2013). The case series by Widdowson and Benelli have shown that TA can be an effective therapy for major depressive disorder when delivered in routine clinical practice, in private practice settings, with clients with mild to moderate impairment in functioning who actively sought out TA therapy and with white British and Italian therapist and client dyads.

The present study analyses the treatment of Anna, a 33-year-old Italian woman who had been suffering from depressive symptoms for several years, worsening in the last few months. Approximately 3% to 6% of all adults in Western countries suffer from a form of depression that persists for at least two years during their lifetime (Kessler, Berglund, Demler, Jin & Merikangas, 2005). The Diagnostic & Statistical Manual of Mental Disorders 5th Edition (DSM-5) (American Psychiatric Association, 2013) has introduced a new diagnostic category of persistent depressive disorder (PDD) that includes the first two of the following four subtypes of persistent forms of depression: (a) a continuing mild depressive mood (dysthymia); (b) a state meeting all criteria for major depression continuously (chronic major depression); (c) a recurrent major depression with incomplete remission between episodes; and (d) a superimposition of a major depressive episode on an antecedent dysthymia (double depression) (Klein, 2010).

The aim of this study was to investigate the effectiveness of the manualised TA treatment of depression (Widdowson, 2016) applied to a persistent depressive disorder (dysthymia). The primary target of the therapy was the depressive symptomatology, the secondary target symptoms were anxiety, global distress and severity of personality problems. Qualitative data was also collected from therapist and client on helpful aspects of the therapy and following change.

Ethical Considerations

The research protocol follows the requirements of the ethical code for Research in Psychotherapy of the Italian Association of Psychology (AIP, 2015), and the American Psychological Association guidelines on the "rights and confidentiality of research participants" (APA, 2010, p. 16). The research protocol has been approved by the Ethical Committee of the University of Padua. Before entering the treatment, the client received an information pack, including a detailed description of the research protocol, and gave an informed consent and written permission to include segments of disguised transcripts of sessions or interviews within scientific articles or for these to be presented at conferences. The client was informed that she would have received the therapy even if she decided not to participate in the research and that she was able to withdraw from the study at any moment without any negative impact on her therapy. All aspects of the case material were disguised, so that neither the client nor third parties are identifiable. All changes are made in such a way as to not lead the reader to draw false conclusions related to the described phenomena. The final article, in Italian language, was presented to the client, who confirmed that it was a true and accurate record of the therapy and gave her final written consent for its publication.

Method

Inclusion and exclusion criteria

Psychotherapists participating in this case series were invited to include in their studies the first new client, with a disorder within the depressive spectrum as described in DSM-5 (Major, Persistent or Other Depressive Disorder), who agreed to participate in the research. Other current psychotherapy, active psychosis, domestic violence, bipolar disorder, antidepressant medication, alcohol or drug abuse were considered as exclusion criteria. As the overall aim of this project is to study the effectiveness of TA psychotherapy in routine clinical practice, comorbidity is normally accepted and both inclusion and exclusion criteria are evaluated case by case.

Client

Anna is a 33-year-old white Italian woman who lives alone in a large metropolitan area in Italy. She is a manager in a tour operator company and loves her job, she reports having a good relationship with her mother, who she described as “an angel” but has a difficult relationship with her father, who has been unable to demonstrate his affection for her in many life circumstances. She reports that her parents have been unable to protect her in some life decisions and situations. For this reason, nowadays she often appears to be very angry with them, especially for not understanding her feelings. She described that in her family everyone over-estimates her capacities to manage everything on her own. She has a younger sister and her parents pressed her to help her sister in finding a job, but

she was not able to secure employment for her. The unemployment of her sister exacerbated her conflictual relationship with her family. In the past years, Anna described feeling responsible for her sister's situation, feeling guilty for achieving and having success in her life, and also feeling culpable for being incapable of doing more to help her sister. Anna felt lonely for her many bad relationship break-ups, which made her think, in the last several years, that there might be something wrong with her manner of relating with men. At the time of therapy, she was single. Anna reported often feeling extremely vulnerable, with periods of intense crying and stomach-aches, and that in the last few years she had some difficulties in falling asleep. She stated she also feels anxious and disappointed in her relationships, and does not get the feedback she expects from her partners, which causes conflict and often in turn leads to men breaking-up with her. She is worried about her future, believes that she will not be able to create a family of her own, and feels that she is not important to anyone. She independently decided to seek therapy and asked a colleague to recommend a therapist.

Therapist

The psychotherapist is a 40-year-old white Italian woman with 8 years of clinical experience and who has a certification as Certified Transactional Analyst (Psychotherapy) (CTA-P). For this case, she received monthly supervision by a Teaching & Supervising Transactional Analyst (Psychotherapy) (TSTA-P) with 30 years of experience.

Intake sessions

Since the client had difficulties in paying for the therapy, the therapist proposed that Anna participated in the research protocol to access lower cost therapy. The client attended four pre-treatment sessions (0A, 0B, 0C, 0D), which were focused on explaining the research project, obtaining consensus, conducting a diagnostic evaluation according to DSM-5 criteria, developing a case formulation and a treatment plan, defining the problems she was seeking help for in therapy, as well as their duration and their severity (i.e. preparing the personal questionnaire, see later), and collecting a stable baseline of self-reported measures for primary (depression) and secondary (anxiety, global distress, personal problems) symtoms.

DSM-5 Diagnosis

During the diagnostic phase, Anna was assessed as meeting DSM-5 diagnostic criteria of mild persistent depressive disorder with mild anxious distress: she experienced depressed mood for more than two years (criterion A) insomnia (B2), reduced self-esteem (B4) and feelings of hopelessness (B6); she also felt excessively anxious (1) and worried (2). Knowing the level of an individual's personality functioning and personality traits provides the therapist with fundamental information for treatment planning. Therefore, a personality diagnosis using the alternative dimensional model developed for DSM-5 Section III was also conducted. This diagnosis allows for assessment of: 1) the level of impairment in personality functioning, and 2) personality traits. Anna showed impairment ranging from little to some in the level of organisation, and personality traits of depressivity, anxiousness, submissiveness, impulsivity, hostility and withdrawal. The therapist also rated the computerised Shedler-Westen Assessment Procedure (SWAP-200) (Shedler & Westen, 1999) that supported the diagnosis of high level of functioning, with traits of depressive, histrionic and dependent personality types.

TA Diagnosis and Case formulation

Anna presented with Be Strong and Please Me drivers (Kahler, 1975) and the injunctions (Goulding & Goulding, 1976) 'Don't be important', 'Don't think', 'Don't be close', and 'Don't be yourself' (Don't be feminine). Anna's racket system (Erskine & Zalcman, 1979) showed beliefs such as 'Be compliant in order to obtain love'. Her script analysis involved substitute feelings (English, 1971) of sadness and anger, with somatisation as defense mechanisms. Interpersonally, Anna tended to alternate dramatic roles (Karpman, 1968) of Victim (when backing down without expressing her feelings), Rescuer (worrying about others, especially her sister), and Persecutor (during outbursts of hostility). Her life position (Ernst, 1971) was I'm Not OK, You're OK.

Treatment

The therapy followed the manualised therapy protocol of Widdowson (2016). The treatment plan focused on creating a therapeutic alliance, primarily providing permission (Crossman, 1966) congruent with the client's injunctions, namely: *you can be important, think, be close, be yourself (feminine)*. The therapist offered Anna empathic listening, supporting her to feel and express her emotions, needs and wishes. During first sessions, the therapist also explained the ego state model, in order to give Anna some theoretical knowledge that might help her to better understand the emotional states she was experiencing and her behaviours. Then, the therapist focused on reinforcing self-esteem, supporting Anna's recognition of the importance of understanding her Child ego state needs for attention and being loved, exploring her experiences, and analysing her script (Steiner, 1966) events such as the relationship with her father, which influences her actual relationships with men.

Analysis Team

The HSCED main investigator and first author of this paper is a Provisional Teaching and Supervising Transactional Analyst (Psychotherapy) (PTSTA-P) with 10 years of clinical experience, with a strong allegiance for TA. Despite recent literature suggesting that hermeneutic analysis should be carried out only by expert psychotherapists (Wall, Kwee, Hu & McDonald, 2016), we decided that when the research is investigating a new population or a therapy that lacks a research base, it is appropriate to follow Bohart (2000), who proposed that analyses can be carried out by a team of 'reasonable persons', not yet overly committed to any theoretical approach or professional role. The team comprised six postgraduate psychology students who were taught the principles of hermeneutic analysis by

Professor John McLeod, in a course on case study research at the University of Padua. Following the indication of Elliott, Partyka, Wagner et al (2009), the students preferred to assume both affirmative and sceptic positions, and independently prepared their affirmative and sceptic cases. Then they met and merged their own cases, supervised by the main investigator, creating a consensual affirmative and sceptic brief and rebuttals.

Transparency statement

The research was conducted entirely independently of the previous case series (see Widdowson 2012a, 2012b, 2012c). The last author, Mark Widdowson, was involved in checking that the research protocol and data analysis process was adhered to, in order to make the claim that this case series represents a valid replication of the initial study (with minor changes) and he was involved in the final preparations of this article.

Judges

The judges were three researchers in psychotherapy at the University of Padua and co-authors of this paper: Judge A, Vincenzo Calvo, clinical psychologist, psychotherapist trained in dynamic psychotherapy, PhD in development psychology, with expertise in attachment theory; Judge B, Stefania Mannarini, psychologist with experience in research methodology; and Judge C, Arianna Palmieri, neuropsychologist and psychotherapist with a training in dynamic psychotherapy. Judges A and C had some basic knowledge of TA but had never engaged in any official TA training, whereas Judge B has some clinical experience but no knowledge of TA.

Quantitative Outcome Measures

Three standardised self-report outcome measures were selected to measure primary target symptoms (depression) and secondary symptoms (anxiety and global distress).

Patient Health Questionnaire 9-item for depression (PHQ-9; Spitzer, Kroenke & Williams, 1999), which scores each of the nine DSM-5 criteria from 0 (not at all) to 3 (nearly every day), which has been validated for use in primary care (Cameron, Crawford, Lawton, et al, 2008). Total scores up to 4 are considered healthy, scores of 5, 10, 15 and 20 are taken respectively as the cut-off points for mild, moderate, moderately severe and severe depression. PHQ-9 score ≥10 has a sensitivity of 88% and a specificity of 88% for major depression (Kroenke, Spitzer, & Williams, 2001) and scores of <10 are considered subclinical.

Generalized Anxiety Disorder 7-item for anxiety (*GAD-7;* Spitzer, Kroenke, Williams, & Löwe, 2006), which scores each of the seven DSM-5 criteria as 0 (not at all), 1 (several days), 2 (more than half the days), and 3 (nearly every day). Total scores of up to 4 are considered healthy, scores of 5, 10, and 15 are taken as the cut-off points for mild, moderate and severe anxiety respectively. Using the threshold score of 10, the GAD-7 has a sensitivity of 89% and a specificity of 82% for GAD and scores of <10 are considered subclinical. It is moderately good at screening three other common anxiety disorders - panic disorder (sensitivity 74%, specificity 81%), social anxiety disorder (sensitivity 72%, specificity 80%) and post-traumatic stress disorder (sensitivity 66%, specificity 81%) (Kroenke, Spitzer, Williams, et al, 2007).

Clinical Outcome for Routine Evaluation - Outcome Measure for global distress (CORE-OM) (Evans, Connell, Barkham, Margison, Mellor-Clark, McGrath, & Audin, 2002). Each of the 34 items is scored on a 5-point scale ranging from 0-4 (0 = not at all, 4 = most of the time). Total scores up to 5 are considered healthy, scores between 5 and up to 9 are considered low level (sub-clinical), and scores of 10, 15, 20 and 25 are taken as the cut-off point for mild, moderate, moderately severe and severe distress, respectively. The cut-off of 10 yields a sensitivity (true positive rate) of 87% and a specificity (true negative rate) of 88% for discriminating between members of the clinical and general populations. CORE-OM was used in assessment sessions, in sessions 8, 16 and follow ups, whereas CORE short form A and B were used in all other sessions (Barkham, Margison, Leach, Lucock, Mellor-Clark, Evans, McGrath et al, 2001).

All measures were evaluated according to Reliable and Clinical Significant Improvement (RCSI) (Jacobson & Truax, 1991). It is important to consider that even under the cut-off score there may be a subclinical disorder. To minimise Type I error (which occurs when cases with no meaningful symptom change are assumed to have improved) we employed also Reliable Change (RC) (Jacobson and Truax, 1991) to evaluate whether observed changes on a measure were statistically reliable and not due to chance. For example, Richards and Borglin (2011) proposed that a minimum reduction of 6 points in the PHQ-9 would be indicative of reliable improvement. Transition from clinical to non-clinical population and reliable change combine to produce a Reliable and Clinically Significant Change Index (RCSI), as robust evidence of recovery in psychological interventions (Evans, Margison & Barkham, 1998; Delgadillo, McMillan, Leach, Lucock, Gilbody & Wood, 2012).

See Table 1 for Clinical Significance (CS) and Reliable Change (RC) values for each employed measure. All these measures were administered prior to the beginning of each session to measure the on-going process and to facilitate the identification of events in therapy that produced significant change.

Before each session, the client also rated the Personal Questionnaire (PQ) (Elliott, Shapiro, & Mack, 1999), a client-generated measure in which clients specify the problems they would like to address in their therapy and rate their problems according to how distressing they are finding each problem (1, not at all; 7, maximum possible). Scores up to 3 are considered subclinical. In this case series, for the PQ we adopted a more conservative RC of two points, rather than the RC of one point already used in the previous case series.

All of these measures were administered in the pre-treatment phase in order to obtain a three-point baseline, and during the three follow-ups, except that in this case Anna's PQ score was not obtained from session 1.

Qualitative Outcome Measurement

The client was interviewed using the Change Interview protocol (CI) (Elliott, Slatick & Urman, 2001) one month after the conclusion of the therapy. The CI is a semi-structured qualitative change measure which asks clients how they feel they have changed during the therapy and how they think these changes came about, what they felt was helpful or hindering in the therapy, and what changes they feel they still need to make. Clients are asked to identify key changes they made and to indicate on a five-point scale: 1) if they expected to change (1=expected; 5=surprising); 2) how likely these changes would have been without therapy (1=unlikely; 5=likely), and 3) how important they feel these changes to be (1=slightly; 5=extremely).

The client also completed the Helpful Aspects of Therapy form (HAT) (Llewelyn, 1988) at the end of each session. The HAT allows the client to describe hindering or useful aspects of the session and to rate them on a nine-point scale (1=extremely hindering, 9=extremely useful).

Therapist Notes

A structured session notes form (Widdowson, 2012a, Appendix 6, p. 50-52) was completed by the therapist at the end of each session. In this form, the therapist provides a brief description of the session in which are identified key aspects of the therapy process, the theories and interventions used, and an indication of how helpful the therapist felt the session was for the client.

Adherence

The therapist, the supervisor, and the main researcher were all Transactional Analysts and they each independently evaluated the therapist's adherence to TA treatment of depression using the operationalised adherence checklist proposed by Widdowson (2012a, Appendix 7, p. 53-55) before agreeing on a final consensus rating. The conclusion of the three evaluators was that the treatment had been conducted coherently according to TA theory at a good to excellent level of application.

HSCED Analysis Procedure

Affirmative Case

The affirmative position according to Elliott (2002) should locate evidence in the rich case record supporting the claim that the client has changed, and that the change is causally due to the therapy. A clear argument supporting the link between change and treatment must be established on the basis of at least two of the following five sources of evidence:

1. Changes in stable problems: client experiences changes in long-standing problems. The change should be replicated in both quantitative and qualitative measures. Change should be Clinically Significant (scores fall in the healthy range), Reliable (corrected for measure error) and Global (Reliable Change is replicated in at least two out of three measures);
2. Retrospective attribution: according to the client the changes are due to the therapy;
3. Outcome to process mapping: refers to the content of the post-therapy qualitative or quantitative changes that plausibly match specific events, aspects, or processes within therapy;
4. Event-shift sequences: links between client reliable gains in the PQ scores and significant within therapy events;
5. Within therapy process-outcome correlation: the correlation between the application of therapy principles (e.g. a measure of the adherence) and the variation in quantitative weekly measures of client's problem (e.g. PQ score).

Sceptic Case

A sceptic position requires a good-faith effort to find non-therapeutic processes that could account for an observed or reported client change. Elliott (2002) identified eight alternative explanations that the sceptic position may consider: four non-change explanations and four non-therapy explanations.

The four non-change explanations assume that change is really not present, and should consider:

1. Trivial or negative change which verifies the absence of a clear statement of change within qualitative outcome data (e.g. CI), and the absence of clinical significance and/or reliable change in quantitative outcome measures (e.g. PHQ9);
2. Statistical artefacts that analyse whether change is due to statistical error, such as measurement error, regression to the mean or experiment-wise error;
3. Relational artefacts that analyse whether change reflects attempts to please the therapist or the researcher;
4. Expectancy artefacts, analysing whether change reflects stereotyped expectations of therapy.

The four non-therapy explanations assume that the change is present, but is not due to the therapy, and should consider:

5. Self-correction which analyses whether change is due to self-help and/or self-limiting easing of a temporary problem or a return to baseline functioning;
6. Extra-therapy events that verify influences on change such as those due to a new relationship, work, or financial conditions;
7. Psychobiological causes which verify whether change is due to factors such as medication, herbal remedies, or recovery from medical illness;
8. Reactive effects of research, analysing the effect of change due to participating in research, such as generosity or goodwill towards the therapist.

The formulation of affirmative and sceptic interpretations of the case consists of a dialectical process, in which affirmative rebuttals to the sceptic position are constructed, along with sceptic rebuttals of the affirmative position.

Finally, each position is summarised in a narrative that offers a customised model of the change process that has been inferred, including therapeutic elements and an account of the chain of events from cause (therapy) to effect (outcome), including mediator and moderator variables.

Adjudication Procedure

Each single judge received the rich case record (session transcriptions, therapist and supervisor adherence forms and session notes, quantitative and qualitative data and also a transcript of the Change Interview) as well as the affirmative and sceptic cases and rebuttals by email, together with instructions. The judges were asked to examine the evidence and provide their verdict. They were required to establish:

- If the case were a clearly good outcome case, a mixed outcome case, or a poor outcome case;
- If the client had changed;
- To what extent these changes had been due to the therapy;
- Which aspects of the affirmative and sceptic arguments had informed their positions.

Furthermore, the judges had to observe which mediator factors in the therapy they considered to have been helpful and which characteristics about the client did they think had contributed to the changes as moderator factor(s).

Results

In earlier published HSCED's the rich case records, along with hermeneutic analysis and judges' opinions were often provided as online appendices (Benelli et al, 2015). Since all the material is in Italian language, we adopted here the solution of providing a summary of the main points, as proposed in MacLeod, Elliott and Rodger (2012). The complete material (session transcriptions, Change Interview, affirmative and sceptic briefs and rebuttal, judge opinions and comments) is available from the first author on request.

Quantitative Outcome Data

Anna's quantitative outcome data are presented in Table 1. The initial PHQ-9 score of 11 indicated a moderate level of depression. The GAD-7 score of 8.3 indicated a subclinical, mild level of anxiety. The CORE at 16.8 indicated a moderate level of global distress and functional impairment. The PQ at 6.2 indicated that the client perceived her problems as bothering her more than very considerably.

At session 8, (mid-therapy), all measures decreased. Depression passed into the subclinical mild range (6), anxiety remained in the mild range (6), global distress passed to subclinical range, with clinically significant and reliable improvement (7.4), and personal problems decreased to moderately bothering (4.3).

By the end of the therapy, the depressive score remained in the mild range (7), the anxiety reliably decreased to healthy range (4), the global distress returned within the mild range (12.1) with a lower score than pre-therapy, and the personal problems reliably decreased (3.7).

At the 1-month follow up, all measures except anxiety improved: depressive scores remained in the mild range (6), anxiety returned to mild (6), the global distress returned to a subclinical range (6.5) with clinically significant and reliable improvement, and personal problems remained reliably improved at moderately (3.5).

At the 3-month follow up, all measures improved: depression passed into the healthy range (3) with a clinically significant and reliable improvement, anxiety reliably decreased to the healthy range (4), global distress entered the healthy range (5) and personal problems were scored as bothering her only a little (3.2).

At the 6-month follow up all scores worsened: depression remained in the healthy range with clinically significant and reliable improvement compared to the pre-therapy (4); anxiety returned to the mild range, with a slight, non-reliable change compared to pre-therapy (6); global distress returned within the mild range (10.29), with a score lower than at the end of the therapy and reliably improved in respect to the beginning of the therapy; personal problems returned to moderately bothering (3.7), with a reliable improvement compared to the pre-therapy score.

Table 2 shows the 10 problems that the client identified in her PQ at the beginning of the therapy and their duration. Four problems were rated as maximum possible, five very considerably and one moderately bothering. Four problems lasted from more than 10 years, two from 6-10 years, three from 3-5 years and one from 1-2 years. Problems lasting for more than 10 years showed a reliable change at the end of the therapy and at the 6-month follow up (except item 8, anger toward parents). All problems lasting from 1-5 years showed an early reliable change within session 8, and of these, three out of four also showed a clinically significant change.

Problems are related to: self esteem (1, incapable; 3, vulnerable); relationships (5 family; 9 colleagues); symptoms (4, guilty; 6, anxiety; 7, sleep) emotions and inner experience (2, loneliness; 8 and 10, anger; 9, oppressed).

At the end of the therapy and at the 1-month follow up, 9 out of 10 problems showed a reliable change, and 3 of these showed also a clinically significant improvement (guilty, oppression, anger). At the 3-month follow up, all problems showed a reliable change, and 5 of these also a clinically significant change. At the 6-month follow up, all problems showed a reliable change (but 8, anger) and guilt, sleeping and feeling of oppression also showed a clinically significant improvement.

	Pre-Therapy[a]	Session 8 Middle	Session 16 End	1 month FU	3 months FU	6 months FU
PHQ-9	**11** **Moderate**	6 (+) Mild	7 (+) Mild	6 (+) Mild	3 (+) (*) Healthy	4 (+) (*) Healthy
GAD-7	8.3 Mild	6 Mild	4 (*) Healthy	6 Mild	4 (*) Healthy	6 Mild
CORE-OM	**16.8** **Moderate**	7.4(+)(*) Low level	**12.1** **Mild**	6.5(+)(*) Low level	5(+)(*) Healthy	**10.29 (*)** **Mild**
PQ	**6.2**[b] **Very considerably**	**4.3** **Moderately**	**3.7(*)** **Moderately**	**3.5(*)** **Moderately**	**3.2(*)** **Little**	**3.7(*)** **Moderately**

Table 1: Anna's Quantitative Outcome Measure

Note. Values in **bold** are within the clinical range; + indicates clinically significant change (CS). * indicates reliable change (RC). CORE = Clinical Outcomes in Routine Evaluation-Outcome Measure (Evans et al., 2002). PHQ-9 = Patient Health Questionnaire 9-item for depression (Spitzer, Kroenke & Williams, 1999) GAD-7 = Generalized Anxiety Disorder 7-item (Spitzer, Kroenke, Williams, & Löwe, 2006). PQ = Personal Questionnaire (Elliott, Shapiro, & Mack, 1999). FU = follow-up.

Clinical cut-off points: CORE-OM ≥10; PHQ-9 ≥10; GAD-7 ≥10; PQ ≥3. Reliable Change Index values: CORE-OM improvement of five points, PHQ-9 improvement of six points, GAD-7 improvement of four points, PQ improvement of two points.

[a]Mean value of pre-therapy assessment sessions. [b]First available score in session 2.

Figures 1 to 4 allow visual inspection of the time series of the weekly scores of primary (PHQ9) and secondary (GAD-7, CORE and PQ) outcome measures, with linear trendline

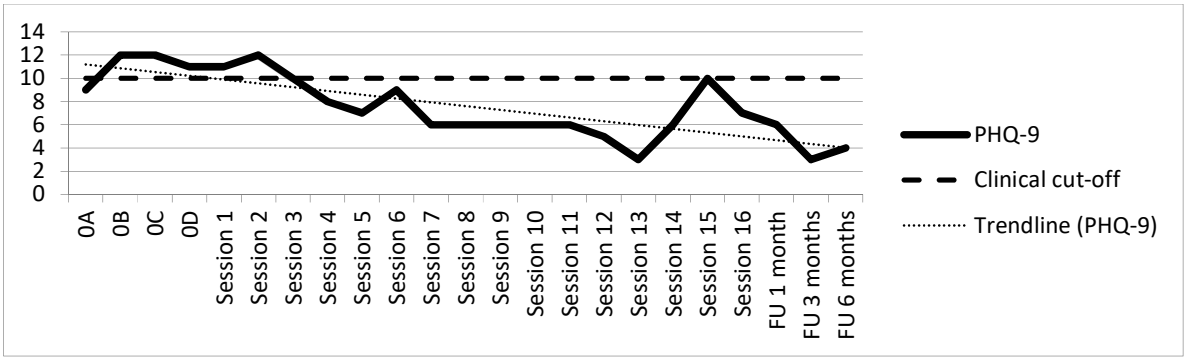

Figure 1: Anna's weekly depressive (PHQ-9) score

Note. 0A, 0B, 0C and 0D = assessment sessions. PHQ-9 = Patient Health Questionnaire 9-item for depression (Spitzer, Kroenke & Williams, 1999). FU = follow-up.

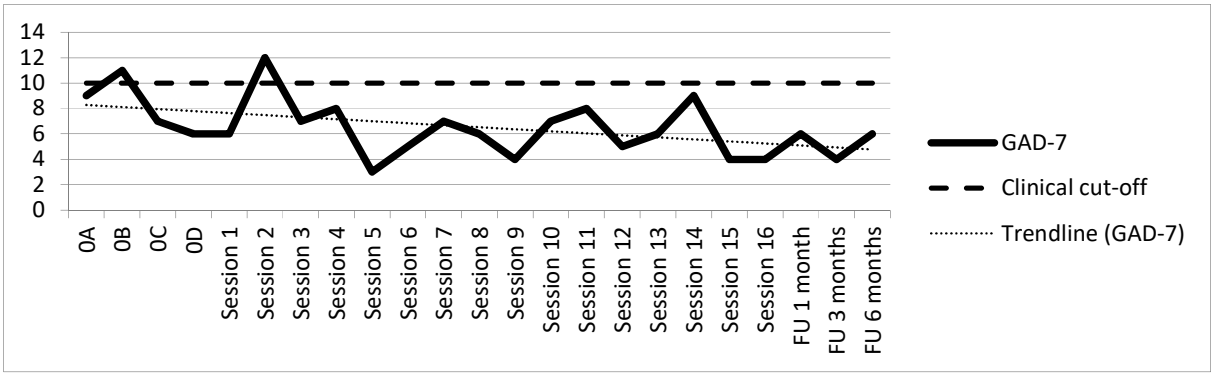

Figure 2: Anna's weekly anxiety (GAD-7) score

Note. 0A, 0B, 0C and 0D = assessment sessions. GAD-7 = Generalized Anxiety Disorder 7-item (Spitzer, Kroenke, Williams, & Löwe, 2006). FU = follow-up.

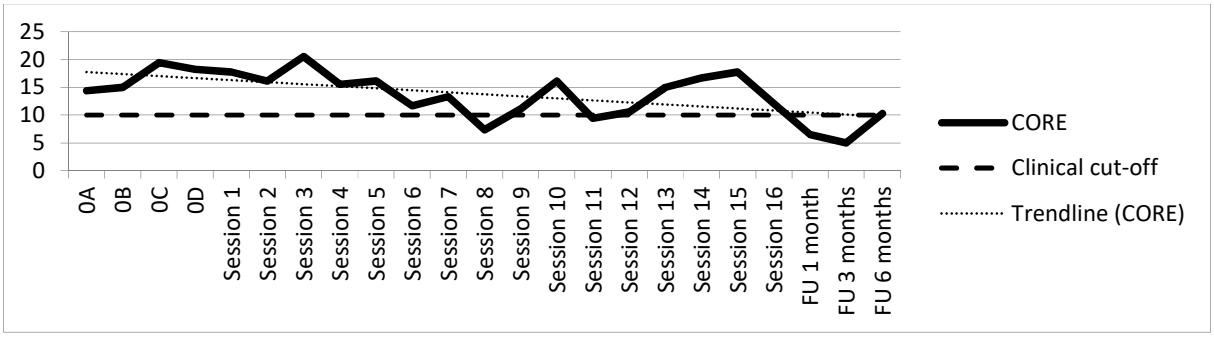

Figure 3: Anna's weekly global distress (CORE) score

Note. 0A, 0B, 0C and 0D = assessment sessions. CORE = Clinical Outcomes in Routine Evaluation-Outcome Measure (Evans et al., 2002). FU = follow-up.

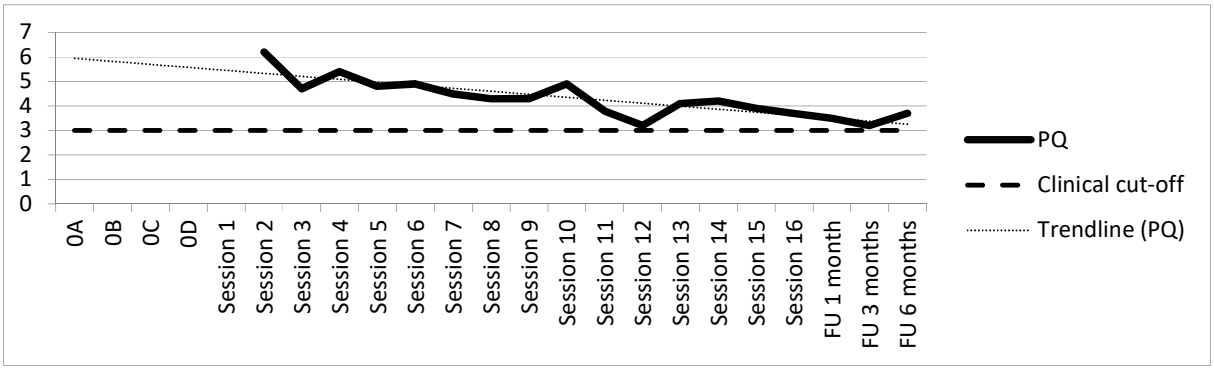

Figure 4: Anna's weekly personal problems (PQ) score

Note. The first available score was in session 2. 0A, 0B, 0C and 0D = assessment sessions. PQ = Personal Questionnaire (Elliott, Shapiro, & Mack, 1999). FU = follow-up.

Qualitative Data

Anna compiled the HAT form at the end of every session (Table 3), reporting only positive/helpful events. All positive events were rated from 7 (moderately helpful) to 9 (extremely helpful). She reported helpful aspects on *self esteem* (HAT 1, appreciate myself; HAT 5, accept myself; HAT 8, faith in myself; HAT 12, reassuring myself); *relationships* (HAT 1, put boundaries; HAT 3, collect information, no expectations; HAT 9, time to understand; HAT 14, receive vs show off); *symptoms* (HAT 2, too responsible); *emotions and inner experience* (HAT 4, utter emotions, become aware; HAT 5, focus feelings; HAT 7 put out anger, awareness of feelings; HAT 8, confidence in myself; HAT 12, stop and think about feelings).

Anna participated in a Change Interview 1-month after the conclusion of the therapy. In this interview, she identified her main and significant changes (Table 4). Anna described her therapy as "very helpful" (Client line 31), "it helped me focus on how to protect myself" (C32). When Anna started the therapy, she felt "more vulnerable" (C35), whereas now she is not "throwing herself headlong in relationships", in fact she is "trying to create a more mature and equal way to relate with men" (C37). At the beginning of the therapy, she felt angry when thinking about people from her past, whereas now she is able to "distance them from my life" (C82). Before starting the therapy, Anna reported feeling guilty for the unemployment of her sister, whereas now she does not feel responsible any more: "Earlier I was focused on my guilt… I changed perspective… I don't feel guilty anymore… I only tried to help her… I did it in good faith" (C92).

Anna summarised four main areas of change. First, she observed an improvement in her way of relating with men. Anna stated that she expected such results, in fact that was her therapy goal (rated 2, somewhat expected), although that she believed that this outcome would have been unlikely without therapy (1) and was very important for her (4). The second change she identified was focusing the aspects she has to work on to improve her affective vulnerability, which she identified as somewhat expected (2) and that the change would have been unlikely to have happened (1) without therapy, rating it as extremely important for her (5). The third and fourth improvements were accepting her past and feeling calmer regarding familiar stress, both rated as neither expected nor unexpected (3), which would have been unlikely to have happened without the therapy (1) and considered respectively as extremely (5) and very

	PQ items	Duration	Pre-Therapy[a]	Session 8 (middle)	Session 16 (end)	1 month FU	3 months FU	6 months FU
1	I feel incapable to develop relationships	>10y	**6 Very considerably**	**5 Considerably**	**4(*) Moderately**	**4(*) Moderately**	3(+) (*) Little	**4(*) Moderately**
2	Deep feeling of loneliness during the weekend	3-5y	**7 Maximum possible**	3(+) (*) Little	**4(*) Moderately**	**4(*) Moderately**	3(+) (*) Little	**5(*) Considerably**
3	I feel affectively vulnerable	>10y	**7 Maximum possible**	**6 Very considerably**	**5(*) Considerably**	**5(*) Considerably**	**4(*) Moderately**	**4(*) Moderately**
4	I feel guilty for my brother's not successful working	6-10y	**7 Maximum possible**	**4**(*) Moderately	3(+) (*) Little	2(+) (*) Very little	2(+) (*) Very little	3(+) (*) Little
5	I feel the familiar stress on me	>10y	**7 Maximum possible**	**5(*) Considerably**	**5(*) Considerably**	**4(*) Moderately**	**5(*) Considerably**	**4(*) Moderately**
6	I feel anxiety for the future	3-5y	**6 Very considerably**	**4(*) Moderately**	**4(*) Moderately**	**4(*) Moderately**	**4(*) Moderately**	**4(*) Moderately**
7	I have difficulties in falling asleep	6-10y	**6 Very considerably**	**6 Very considerably**	**5 Considerably**	**4(*) Moderately**	**4(*) Moderately**	3(+) (*) Little
8	I feel angry for being left alone at school	>10y	**6 Very considerably**	**6 Very considerably**	**4(*) Moderately**	**5 Considerably**	**4(*) Moderately**	**5 Considerably**
9	I feel oppressed by a colleague's presence	1-2y	**4 Moderately**	1(+) (*) Not at all	1(+) (*) Not at all	1(+) (*) Not at all	1((+)*) Not at all	1(+) (*) Not at all
10	I feel anger when thinking about my exes	3-5y	**6 Very considerably**	3(+) (*) Little	2(+) (*) Very little	2(+) (*) Very little	2(+) (*) Very little	**4(*) Moderately**
	Total		62	43	37	35	32	37
	Mean		**6.2 Very considerably**	**4.3 Moderately**	**3.7(*) Moderately**	**3.5(*) Moderately**	**3.2(*) Little**	**3.7(*) Moderately**

Table 2: Anna's personal problems (PQ), duration and scores

Note: Values in **bold** are within clinical range. PQ = Personal Questionnaire (Elliott, Shapiro, & Mack, 1999). Clinical cut-off point: PQ ≥3. Reliable Change: PQ improvement of two points. +=indicates clinically significant change (CS). *=indicates reliable change (RC). The rating is on a scale from 1 to 7 and indicate how much each problem has bothered the client: 1 = not at all; 7 = maximum. m = months. y = year. FU= follow-up.

[a]The first available score was in session 2.

.

Session	Rating	Events	What made this event helpful/important
1	8 (greatly)	From the dialog with the therapist emerged that I have to learn to appreciate myself more and to put boundaries in my relationships with men	It's important because I want to learn to evaluate myself more, to feel desired by others
2	8 (greatly)	Being able to talk and cry about painful events of my family members (parents and brother)	Being able to understand that I feel too responsible for others: a weight too heavy for me
3	8 (greatly)	The session's theme was [PQ] point 1 "I feel incapable to develop relationships". The important aspect is to try to change my approach: do not interpret, but collect information	"Collect information" means do not throw yourself headlong into someone; I don't have to make expectations if there is no feedback on the other side
4	9 (extremely)	Utter my fears while thinking at the person I like and at a possible relationship with him	Utter my fears means to become aware and work more on myself to get better (and not being scared any longer)
5	8 (greatly)	Being capable to better focus what I feel (especially the inappropriateness I feel in some situations)	It's important to focus on my learning to accept myself and not feeling "wrong"
6	-	Missing	Missing
7	8 (greatly)	"Being able to pull out the anger I felt towards my father, since I was a kid"	It helps me to become aware of what I feel, what I need to cure the Child I am
8	8 (greatly)	Talk about my "contract", that is expressing what I can about my self-awareness, regaining faith in my self	It's important to talk about what I'm living right now in order to acquire more confidence in myself
9	8 (greatly)	Give me time to understand a relationship	I learnt that it's important to invest my time
10	7 (moderately)	Gain awareness about a desired relationship without having a positive development	Gaining the awareness
11	8 (greatly)	It has been a very painful session for me. I feel like a disaster when dealing with feelings with the other sex	I feel very lonely
12	8 (greatly)	Trying to understand how to "not hit the ground running" when taking decisions about affective feelings, stop and think in order to be more aware of the choice I made	It has been helpful to understand how to "take by hand" the Child in me, reassuring her from her fears
13	9 (extremely)	Talk about a trauma of the past that caused me pain (being forced to be the only girl in a class of boys, from the fifth to the ninth grade)	I hope talking will help me to heal from that pain
14	8 (greatly)	Being able to talk about my "modus operandi" in working and affective situations, and reveal that in the affective ones I'm always the first one to show off, instead I have to learn to receive and to be seen	Trying to learn to be seen, to receive
15	9 (extremely)	Realize what I want for myself in a relationship	It has been important because I have to change my modus operandi: I don't want to be the only one to give, I want to receive too
16	8 (greatly)	For me, has been helpful to make the point of what I understood of the therapy and what I want for myself and for my future	It's helpful to speak about what I feel, to be the main focus of attention: not being the only one to give, but to receive too

Table 3: Anna's helpful aspect of therapy (HAT forms)

Note. The rating is on a scale from 1 to 9: 1 = extremely hindering, 5 = neutral, 9 = extremely helpful. HAT = Helpful Aspect of Therapy (Llewelyn, 1988).

Change	How much expected change was [A]	How likely change would have been without therapy [B]	Importance of change [C]
I feel capable to develop relationships	2 (somewhat expected)	1 (unlikely)	4 (very)
I focus the aspects to work on (my affective vulnerability) due to the *modus operandi* I had with my father	2 (somewhat expected)	1 (unlikely)	5 (extremely)
Improved relationship with my family, learned to accept the past	3 (neither)	1 (unlikely)	5 (extremely)
Calmer relative to familiar stress	3 (neither)	1 (unlikely)	4 (very)

Table 4: Anna's Changes identified In the Change Interview

Note. CI = Change Interview (Elliott et al., 2001).

[a]The rating is on a scale from 1 to 5; 1= expected, 3 = neither, 5 = surprising. [b]The rating is on a scale from 1 to 5; 1=unlikely, 3 = neither, 5 = likely. [c]The rating is on a scale from 1 to 5; 1 = slightly, 3 = moderately, 5 = extremely.

important. Anna also reported that a friend of hers realised that she was calmer (C41). "Once the therapist told me that it's ok to hit the rock, but afterwards you can raise back up… I feel better knowing this, because there were days in which I felt like a total failure" (C49). Anna in her CI did not report any negative, obstructive or unpleasant aspect of therapy. On the contrary, she felt that some sessions were "really painful, yet crucial to focus on my problems… I metabolised the pain… I have a stronger will to feel better" (C67-8).

HSCED Analysis

Affirmative Case

The affirmative team identified four lines of evidence supporting the claim that Anna changed and that the therapy had a causal role in this change.

Change in stable problems

Quantitative data (Table 1) shows that there is an early improvement in primary outcome measure (depression) that is clinically significant since session 8 and with reliable and clinically significant improvement (RCSI) at 3- and 6-month follow up. Secondary outcome measures depict a reliable improvement in the initial subclinical score of anxiety (GAD-7) at the end of the therapy and at the 3-month follow-up. There is also an early change with RCSI for global distress (CORE) at session 8, maintained at 1- and 3- month follow up. In the PQ (Table 2), Anna identified 10 main problems at the beginning of the therapy that she was trying to solve, almost all rated as bothering her very considerably (6) to maximum possible (7), nine out of 10 standing from 3 to more than 10 years. All the problems referred to issues with self esteem, relationships, symptoms, emotions and inner experience. At the end of the therapy and at the 6-month follow up 9 problems out of 10 showed reliable change, and three problems also reached RCSI. Overall, there is support for a claim of global reliable change (reliable change in at least two out of three measures). Qualitative data supports this conclusion: in fact, in her Change Interview (CI) Anna reports as a main achievement in therapy her change in dealing with others, men, family and her past experiences, all problems rated in the PQ as long standing (more than 10 years). At the end of the therapy she also appears more capable of asserting herself (session 15, C33-35), that implies a change in self experience (vulnerable), another long bothering problem since more than 10 years. Since sessions 7 (C16-40) and 8 (C5-8) Anna showed up with a higher mood. Thus, we claim that Anna obtained a stable RCSI in persistent depressive disorder, and a reliable improvement in global distress and in long-standing problems.

Retrospective attribution

Anna identified in her Change Interview four important changes in different aspects of her life, all of which she attributed to therapy (Table 4). She considered her improvements very and extremely important, and stated that she believed all were unlikely to have occurred without therapy, with the first two changes expected and the last two neither expected nor unexpected. She recognised that the therapy allowed her to change different aspects of her way of relating with men (CI, C35), which was directly related to her therapy contract. The client asserted that the therapy was very useful to her, in particular for the kind of mature and equal relationships she feels she is now capable of establishing (CI, C35). She also affirmed that there were no negative aspects, obstacles or unhelpful aspects to her therapy. In session 16, Anna reported being sad about the ending of the therapy, because it had helped her to focus on her problems and learn what she needed to work on to change for the better (session 16, C155-156). Due to the

new strategies she had been working on with her therapist and started to use in her everyday life (see Table 3), she had noticed positive changes.

Association between outcome and process (outcome to process mapping)

The HAT completed at the end of each session provides us with regular and immediate reports of what Anna found helpful in each session. All reported events are considered moderately to extremely useful and are coherent with both the diagnosis and the interventions reported in the therapist's notes. One of the client's most important changes reported in the Change Interview refers to the ability to "focus on the aspects I need to work on" (Table 4, CI C31) that appear tied to the therapist's frequent interventions on the importance of Anna clearly and succinctly expressing what she feels and thinks. This is mirrored in the client's HAT 5 ("be able to focus better on what I feel"), 9 ("give me time to understand"), 12 ("stop and think"), 15 ("realise what I want") and 16 ("what I understand, what I want"). Also, the change "I was unable to develop relationships" appears tied to the therapist's interventions reported by the client in the HAT 1 ("to put boundaries in place"), HAT 3 ("do not interpret, but collect information"), 14 ("I show off") and 15 ("realise what I want in relationships"). The other change about family ("accept the past and feel calmer") appears connected to the HAT 2 ("talk and cry about painful family events"), 7 ("pull out the anger towards my father"), 13 ("talk about a trauma of the past").

Event-shift sequences (process to outcome mapping)

The PQ mean score shows a progressive decrease in severity of her problems from the initial score (6.2, more than very considerably) to the final score (3.7, less than moderately). The therapist's confrontation of the client's tendency to feel responsible for others, in particular her sister's employment problems (session 2), is reflected in the PQ item 4 (guilty), that decreased since session 3, achieved RCSI in session 9 and was maintained throughout the entire follow-up period. The interventions regarding her tendency to 'please others' in session 5 led the client to become aware of her anger and to use it for taking an assertive position with her family and colleagues (session 5, C 174). This was reflected in improved scores in PQ item 9 that reached RCSI since session 5 and was maintained at the 6-month follow up.

Sceptic Case

1.The apparent changes are negative (i.e. involved deterioration) or irrelevant (i.e. involve unimportant or trivial variables).

The client entered the trial with moderate depression (PHQ-9, score 11), barely over the threshold for major depressive disorder. Considering the typical cyclical pattern of the diagnosed persistent depressive disorder, it is quite likely that a natural reversal may occur in the following months. Change on anxiety (GAD-7) is irrelevant since the initial score was subclinical and change is not maintained at the 6-month follow up. The global distress score (CORE) shows an inconsistent pattern, and remains in the clinical range at the end of the therapy and at the 6-month follow up. Reliable change is present in three measures out of four, and RCSI is present only for primary outcome, suggesting that a claim of Global Reliable Change is unwarranted. Also in qualitative data, we note evidence of inconsistent change: at session 15, Anna tells about an episode she had with some friends, in which they told her she is not improving in her way of relating to men. Furthermore, at the final session, she reports ruminating on whether she did the right thing with a man she liked. During the CI, the client states that she has not completely worked on her insecurity, and still feels frustrated when dealing with stressful people (like her boss). In the 3-month follow up, Anna still refers to feeling guilty about her sister's unemployment and that she sought explanations from the executive director at her company about why they did not offer her sister a job. Thus, change reported in quantitative self-reported measures does not appear to be supported by the client's statements. Thus, we conclude that the change observed in the primary outcome is more due to the typical pattern of persistent depressive disorder than to the therapy, and only a longer follow-up could determine the effect of the therapy.

2. The apparent changes are due to statistical artefacts or random errors, including measurement error, experiment-wise error from using multiple change measures, or regression to the mean.

The pre-treatment baseline related to the PQ has not been collected due to technical problems, and the score is available only from the second session, making it difficult to draw any conclusions on change in relation to long-standing personal problems, due to missing a stable baseline which would enable clear comparison with subsequent scores. We also noticed that Anna's scores for the PQ item 7 ("I have difficulties in falling asleep") has correspondence to the GAD-7 item 4, in PHQ-9 item 3 and in CORE-B item 2, and they received different evaluation in the same session. For example, at session 13, the client scored this item 2 (sometimes) in the CORE-B, 1 (several days) in the PHQ-9, 2 (over half the days) in the GAD-7, and 5 (considerably) in the PQ. This suggests that the client might have some difficulties in relating her inner experience to scoring of individual items, thus introducing a possible inconsistency within quantitative results.

3. The apparent changes reflect relational artefacts such as global hello-goodbye effects on the part of a client expressing his or her liking for the therapist, wanting to make the therapist feel good, or trying to justify his or her ending therapy.

In her CI, Anna reported only positive comments about the therapy and the therapist, and in her HAT forms she reported only positive/helpful events. Even session 11, after which she "forgot" to attend the following session without informing the therapist, is described as helpful, and the event is not mentioned in the CI. This suggests that CI and HAT may be biased by Anna's tendency to 'please others' and a desire to present a good image of

her therapist to the researcher conducting the CI. Her 'please others' tendency is also in line with her diagnosis. Furthermore, Anna keeps asking the therapist whether she is doing the right thing or not (e.g. in session 14: C2, C45, C63), and comments that she was looking forward to the day of her session in order to talk to the therapist and ask her for advice (session 15, C 73), supporting the conclusion that she tends to depend on the therapist's advice and approval which could have affected her outcome measures.

4. The apparent changes are due to cultural or personal expectancy artefacts; that is, expectations or scripts for change in therapy.

The sceptic team were not able to find any evidence within the rich case record that would support a claim that Anna's changes were associated with expectancy effects.

5. There is credible improvement, but it involves a temporary initial state of distress or dysfunction reverting to normal baseline via corrective or self-limiting processes unrelated to therapy.

DSM-5 indicates that the typical pattern of persistent depressive disorder is likely to include a major depressive episode that may spontaneously revert to a subclinical level. The primary outcome measure could have captured this spontaneous cyclical pattern.

6. There is credible improvement, but it is due to extra-therapy life events, such as changes in relationships or work.

The sceptic team were not able to find any evidence within the rich case record which would support a claim that Anna's changes were associated with extra-therapy life events.

7. There is credible improvement, but it is due to psycho-biological processes, such as psychopharmacological mediations, herbal remedies, or recovery of hormonal balance following biological insult.

In the CI, Anna reports she has not been taking any kind of drugs (T7-C8). The sceptic team were not able to find any evidence within the rich case record which would support a claim that Anna's changes were associated with psychotropic medication or other herbal or similar kind of remedy.

8. There is credible improvement, but it is due to the reactive effects of being in research.

Participating in the research implied a lower cost for the client, and this might have more or less unwittingly affected the rating scores, probably in interaction with the abovementioned 'please others' effect.

Affirmative Rebuttal

Global Reliable Change in the literature is referred to as a measure to control experiment-wise error, thus relying on reliable change and not on clinical significance (e.g., Elliott, 2002). Thus, we can claim that three out of four measures support a claim in favour of Global Reliable Change. Despite the typical cyclical patterns of the per-srsistent depressive disorder, the client identifies change in long-standing problems that were not resolved in previous years by the simple passage of time or natural course of the disorder. Inconsistency between client statements and outcome measures are evidence that stable change is a process achieved during therapy and gradually displayed after its end; the client reports being aware of still having issues to work on (session 15, C156); in the first follow-up the client expressed the desire to continue in the therapy (FU1, C55); and in the third follow-up she complained about not having been contacted to resume the treatment.

Thus, the deterioration observed in quantitative measure at the third follow-up, is not supported by the client's verbal reports on daily life, and may reflect her desire to appear as experiencing greater suffering and therefore be more needing of treatment in order to continue the therapy. Despite missing a clear baseline score for the PQ, we can assume that the score obtained in the second session is representative of the baseline score, since all problems were long standing in time. As for the supposed difficulty of the client in rating self-report measures, the validity of the outcome instruments is widely established and personal scores are corrected for measurement errors by the use of reliable change index. Regarding the 'please others' effect, in the HAT 15, in the CI and in the 1-month follow up, Anna says that some sessions had been painful for her (CI, C67; FU1, C83), showing an ability to critically appraise her own therapy.

During the Change Interview and the follow-ups, the client refers to her sense that sixteen sessions were not sufficient for achieving all of her desired results, and at the 1-month follow up she described herself as satisfied with how she had learnt to handle stressful and painful situations (C37), but asked to continue the therapy (without fee reduction) to allow her to work on more general problems related to her personality, supporting the claim that being included in research did not affect the outcome.

Sceptic Rebuttal

Despite lack of agreement within the literature on how to determine Global Reliable Change, it would be more conservative to claim Global Reliable Change only when *both reliable change and clinical significance are achieved in two out of three measures.* Within transcripts (session 11, 14 and 15) of the therapy it is possible to find evidence that although the client did show some signs of improvement, these were not of a sufficient magnitude to warrant a claim of 'Global Reliable Change of the client. Anna appears to recognise when she is making unfair expectations of men, yet she is still creating fantasies about them. On the other hand, it seems she is now able to feel free from her sense of guilt, but sometimes she acts in order to pacify her guilty feelings. The deterioration at the 6 month follow up suggests that the treatment did not obtain a stable change.

Affirmative Conclusion

Anna's dysthymia, anxiety, global distress and personal problems were related to difficulties in interpersonal patterns, in particular with men, and intrapsychic patterns and inner experience such as emotions and self-esteem. Since the beginning of therapy, the therapist created a climate where the client explored an ability to appreciate herself, expression of emotions such as guilt, sadness and anger, new behaviours such as putting boundaries in place with others, and achieving a new comprehension of her inner experience, thus allowing herself more time to reflect on her emotions and needs before acting. Furthermore, the client explored connections between present and past relational patterns, differentiating past and present. These experiences were reflected in changes in internal dialogues, interpersonal relationships, depressive symptoms, and personality traits of submissiveness, anxiety, hostility and impulsivity.

Sceptic conclusion

Anna asked for therapy during a deterioration in her otherwise subclinical or normal depressive symptomatology. During the therapy, the typical pattern of her persistent depressive disorder reverted to the normal range. Her personality traits (submissiveness, dependent) affected her relationships with the therapist and probably her outcome scores. Changes in intrapsychic and interpersonal patterns are therefore likely to be due to the spontaneous remission of symptoms and to the reassuring effect provided by the presence of the therapist on her personality traits.

Adjudication

Each judge examined the rich case record and hermeneutic analysis and independently prepared their opinions and ratings of the case (Table 5). The judges overall conclusions are that this was a clearly good outcome case, that the client changed substantially, and that the changes are between substantially and completely due to the therapy.

Opinions about the treatment outcome (good, mixed, poor)

Judge A (VC). This case appears to be a clearly good outcome (80% certainty) with some aspect of mixed outcome (20%). Quantitative data shows a reliable and clinically significant change on measures of primary outcome (PHQ-9) at 3- and 6-month follow-ups. Measures of secondary outcomes also improved and there is a Global Reliable Change with three out of four measures showing a reliable change. PQ scores and qualitative data supports the conclusion that a change in long-standing problems occurred: for example, relationships with partners, colleagues and family are fully described as improved.

Judge B (SM). This is a clearly good outcome (80% certainty) or a mixed outcome (20%). The primary outcome was that the client's depressive symptoms passed from moderate into the healthy range during the course of the treatment. The final sessions and the Change Interview report clear descriptions of change in the client's life.

Judge C (AP). This case is classifiable as a good outcome case (80%) to mixed case (20%). Considering quantitative primary and secondary outcomes, every measure improved at the end of the therapy, indicating a change in depression, general distress, anxiety and severity of personal problems.

Opinions about the degree of change

Judge A. The client changed substantially (80% with 80% certainty). Quantitative measures support the claim that the client's PHQ9 shows a stable healthy score at 6-month follow-up, indicating a change in persistent, long-standing depressive symptomatology. The problems reported at the beginning of the therapy in the PQ were almost all long-lasting problems, bothering her up to ten years, and almost all problems show a reliable decrease. In the Change Interview, the client described a clear change in self-representation (guilty vs no longer guilty, vulnerable vs no longer vulnerable) indicating deep changes in personality dimensions and not only symptomatic modifications.

Judge B. The client changed considerably (60% with a 100% certainty), above all in her relationships with others and her family, and reported a decrease in problems described in the PQ. Despite there being some doubt about a claim of Global Reliable Change, the dysthymic symptoms are still absent six months after the end of therapy, indicating a deep and stable change in symptoms and in depressive personality traits. The client reported detailed pre-post differences in relationships with her parents, sister, friends, and colleagues, and a different stance towards her own internal experience.

Judge C. The client showed a substantial change (80% with 80% of certainty) in quantitative and qualitative data. Changes in long standing interpersonal relationships (specifically her sister and parents) support the conclusion that a deep and stable change happened. A longer follow up could further explain the degree and stability of change.

Opinions about the causal role of the therapy in bringing the change

Judge A. The observed change is substantially (80% with 80% of certainty) due to the therapy. Quantitative PQ scores change following interventions that are reported as very important and helpful both in the therapist's notes and in the client's HAT forms. The focus of the therapist on the past experience that still influences the present and the differentiation between present and past appear tied to relational change between sessions, as reported in verbatim transcriptions. Qualitative data (Change Interview) reports clearly retrospective attribution of the client's four main changes to the therapy. HAT forms (summarised in Table 3) are rich in information on what happened during the sessions, and they appear coherent with the change the client feels she has obtained and which she described in the Change Interview.

Judge B. The change is substantially (80% with 80% of certainty) due to the therapy. The client refers to several helpful aspects in her HATs, and clearly states that her main change would have been unlikely without therapy. She reports changes in relationships that appear clearly connected to interventions in psychotherapy (e.g., feeling guilty vs recognising her sister's responsibility for her own situation). In the CI the client clearly defined her changes as unlikely without therapy.

Judge C. The change appears completely due to the therapy (100% with 80% of certainty). Hermeneutic analyses provide a clear link between specific therapeutic foci and changes in PQ scores. It appears unlikely that the client could change her relational patterns without the interventions of the therapist, as described in the HAT forms.

Mediator Factors

Judge A. A good therapeutic alliance and equal relationship appear important for the client's change in therapy. Explanation of the ego state model in the first sessions appears to have been a strong mediator of agreement on goals and alliance. The therapist focused the attention of the client during the sessions, modelling an ordered exploration of events rather than impressionist descriptions and impulsivity. The client's internal dialogue which generated feelings of guilt have been explored, examined, and reappraised. Behavioural submissiveness and a tendency to withdraw were challenged and reappraised. Confrontation of maladaptive patterns, such as feeling guilty for others' failure, allowed change in depressive symptoms and personality traits. Differentiation between here and now and there and then emotional reactions to the stimuli allowed a change in interpersonal patterns.

Judge B. The client-therapist relationship is equal, with the therapist taking an active stance in the therapy, but without leading or suggesting. The therapist paid attention to helping the client to remain focused in the therapy and in defining vague and unspecific statements about events, feeling and behaviours, thus addressing personality traits of withdrawal, impulsivity and submissiveness. Systematic, early exploration of connections between present and past relationships appears tied to enhanced awareness and change in relationships. Focus on self-protection allowed the client to self-explore new behaviours in old relationships.

Judge C. In a manner which was coherent with the diagnosis of don't think and don't be important injunctions, the therapist focused on correspondent permissions, which supported the development of an early alliance. The therapist focused on promoting change in the client's interpersonal behaviours of submissiveness and withdrawal, and in supporting the exploration of alternative behaviours. The therapist focus on the difference between past and present relational experience supported the client in developing insight into attitudes learned in the past which were no longer appropriate in the present.

Moderator Factors

Judge A. The client appears able to immediately assume the 'client role'. She appears motivated, actively seeking therapy, with a high level of personality organisation and intelligence.

Judge B. The therapy was probably enhanced by moderator factors such as: the client's level of higher education, intelligence, and high level of personality functioning.

Judge C. The client was motivated, collaborative and willing to explore her inner world, and open to the therapist's interventions, and was searching for a caring relationship.

Discussion

This case aimed to investigate the effectiveness of a manualised TA treatment for depression in a client with moderate level of persistent depressive disorder (PDD). Primary target was depressive symptomatology, that

	Judge A VC	Judge B SM	Judge C AP	Mean
How would you categorize this case?	Clearly good outcome	Clearly good outcome	Clearly good outcome	Clearly good outcome
How certain are you?	80%	80%	80%	80%
To what extant did the client change over the course of therapy?	80% Substantially	60% Substantially	80% Substantially	73% Considerably to Substantially
How certain are you?	80%	100%	80%	87%
To what extent is this change due to therapy?	80% Substantially	80% Substantially	100% Completely	87% Substantially to Completely
How certain are you?	80%	80%	80%	80%

Table 5: Adjudication results.

showed a reliable change since session 8 and a clinically significant change since the 3-month follow up, maintained in the 6-month follow up. According to DSM-5, the course of PDD show a typical pattern with symptoms rising to the level of a major depressive episode, followed by a likely reversion to a lower level.

Symptoms in PDD are much more unlikely to resolve compared to a Major Depressive Disorder, and thus current clinical practice guidelines recommend the use of psychotherapeutic treatments for PDD (American Psychiatric Association, 2010; NICE, 2009). The therapist conducted the treatment with a good to excellent adherence to the manual. Hermeneutic analysis pointed out changes in stable problem, retrospectively attributed to the psychotherapy, highlighting connections between outcome and process. The judges concluded that this is a clearly good outcome case, with a considerably to substantial degree of change, substantially to completely due to the therapy. These conclusions provide supporting evidence as to the effectiveness of manualised TA psychotherapy for depression, and provide evidence that the manual is suitable for use with persistent depressive disorder.

The therapeutic alliance appears to have been built on a non-directive but active style, focused on personality traits associated to symptoms and addressing their origin in the past. Specific TA techniques were: exploration of internal dialogue, developing the client's Nurturing Parent, exploration of the Be Strong and Please Others' drivers, racket analysis of guilt, sadness and hostility, disconnecting rubberbands (Kupfer & Haimowitz), game analysis (Berne) and analysis of drama triangle roles.

Limitations

The first author has a strong allegiance to TA, is a teacher of the members of the hermeneutic groups and a colleague of the three judges. The author was also funded for this research by TA institutions (see Funding below). Despite the reflective attitude adopted in this work, these factors may have influenced in subtle ways both the hermeneutic analysis and the judges' evaluations.

Conclusion

This case study provides evidence that the specified manualised TA treatment for depression (Widdowson, 2016) has been effective in treating a persistent depressive disorder in an Italian client-therapist dyad. Despite results from a case study being difficult to generalise, this study adds evidence to the growing body of research supporting the efficacy and effectiveness of TA psychotherapy, and notably supports the effectiveness of manualised TA psychotherapy for depression as applied to persistent depressive disorder.

Funding

This study was supported by grants from the European Association for Transactional Analysis, as part of the project 'Transactional Analysis meets Academic Research in order to become an Empirically Supported Treatment: an Italian two-year plan for publishing evidence of Transactional Analysis efficacy and effectiveness into worldwide recognized scientific journals' and from the Center for Dynamic Psychology - Padua, a transactional analysis-oriented School of Specialization in psychotherapy.

Enrico Benelli PhD, Provisional Teaching and Supervising Transactional Analyst (Psychotherapy), Vice-President of the Center for Dynamic Psychology of Padua (Italy), Postdoctoral Researcher, University of Padua, can be contacted at: enrico.benelli@unipd.it

Emanuela Moretti, Certified Transactional Analyst (Psychotherapy), SIFP - Italian Society for Training in Psychotherapy

Giorgio Cristiano Cavallero, Teaching and Supervising Transactional Analyst (Psychotherapy), SIFP - Italian Society for Training in Psychotherapy

Giovanni Greco, Psychotherapist, Founder of SIFP - Italian Society for Training in Psychotherapy

Vincenzo Calvo, Assistant Professor, University of Padua

Stefania Mannarini, Associate Professor, University of Padua

Arianna Palmieri, Assistant Professor, University of Padua

Mark Widdowson, PhD, Teaching & Supervising Transactional Analyst (Psychotherapy), Senior Lecturer, University of Salford

Acknowledgements

Giovanni Greco left us during the final stages of writing of this article, after having contributed to the development and implementation of the research design as Director of the SIFP - Italian Society for Training in Psychotherapy. We thank the new SIFP Director Claudia Filipponi for carrying on his scientific legacy. We thank Mariavittoria Zanchetta for her collaboration in organizing the rich case record and hermeneutic analysis.

References

AIP - Italian Association of Psychology (2015). Ethical code for research in psychotherapy. http://www.aipass.org/node/26 Accessed 10 January 2017

American Psychiatric Association. (2010). *Practice guideline for the treatment of clients with major depressive disorder* (3rd ed.), Arlington, VA: American Psychiatric Association

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

APA - American Psychological Association (2010). Publication Manual (6th edition). Washington DC: Author

Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., Benson, L., Connell, J., Audin, K., McGrath, G., (2001). Service profiling and outcomes benchmarking using the CORE-OM: Toward practice-based

evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, Vol 69(2), 184-196.

Benelli, E., De Carlo, A., Biffi, D., & McLeod, J. (2015). Hermeneutic Single Case Efficacy Design: A systematic review of published research and current standards. *Testing, Psychometrics, Methodology in Applied Psychology, 22*, 97-133. DOI:10.4473/TPM22.1.7

Benelli, E, Revello, B, Piccirillo, C, Mazzetti, M, Calvo, V, Palmieri, A, Sambin, M and Widdowson, M. (2016a), TA treatment of depression: a hermeneutic single-case efficacy design study - 'Sara'. *International Journal of Transactional Analysis Research*, 7 (1) , pp. 3-18.

Benelli, E, Scotta, F, Barreca, S, Palmieri, A, Calvo, C, de Renoche, G, Colussi, S, Sambin, M and Widdowson, M. (2016b). TA treatment of depression: a hermeneutic single-case efficacy design study - 'Penelope'. *International Journal of Transactional Analysis Research*, 7 (1) , pp. 19-34.

Benelli, E, Boschetti, D, Piccirillo, C, Quagliotti, L, Calvo, V, Palmieri, A, Sambin, M and Widdowson, M. (2016c). TA treatment of depression: a hermeneutic single-case efficacy design study - 'Luisa'', *International Journal of Transactional Analysis Research*, 7 (1) , pp. 35-50.

Bohart, A. C. (2000). *A qualitative "adjudicational" model for assessing psychotherapy outcome*. Paper presented at the meeting of the Society for Psychotherapy Research, Chicago.

Cameron, I. M., Crawford, J. R., Lawton, K., et al. (2008). Psychometric comparison of PHQ-9 and HADS for measuring depression severity in primary care. *Br J Gen Pract.*;58(546):32-6. DOI: 10.3399/bjgp08X263794

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of consulting and clinical psychology, 66*:1, 7-18.

Crossman, P. (1966). Permission and Protection. *Transactional Analysis Bulletin, 5*, 152-4.

Delgadillo, J., McMillan, D., Leach, C., Lucock, M., Gilbody, S., & Wood, N. (2014). Benchmarking routine psychological services: a discussion of challenges and methods. *Behavioural and cognitive psychotherapy*, *42*(01), 16-30.

Elliott, R. (2002). Hermeneutic Single-Case Efficacy Design. *Psychotherapy Research*, *12*(1), 1-21. DOI:10.1080/713869614

Elliott, R., Partyka, R., Wagner, J., Alperin, R., Dobrenski, R., Messer, S. B., Watson, J.C., & Castonguay, L. G. (2009). An adjudicated hermeneutic single-case efficacy design study of experiential therapy for panic/phobia. *Psychotherapy Research, 19*(4-5), 543-557. DOI:10.1080/10503300902905947

Elliott, R., Shapiro, D. A., & Mack, C. (1999). *Simplified Personal Questionnaire procedure manual*. Toledo, OH: University of Toledo.

Elliott, R., Slatick, E., & Urman, M. (2001). Qualitative change process research on psychotherapy: Alternative strategies. *Psychologische Beiträge, 43*, 69-111.

English, F. (1971). The substitution factor: rackets and real feelings. *Transactional Analysis Journal, 1*, 4, 225-230).

Ernst, F. H., Jr. (1971). The OK corral: The grid for get-on-with. *Transactional Analysis Journal,* 1(4), 33-42.

Erskine, R. & Zalcman, M. (1979). The racket system: a model for racket analysis. *Transactional Analysis Journal, 9*, 51-9.

Evans, C, Connell, J., Barkham, M., Margison, F., Mellor-Clark, J., McGrath, G. & Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry*, 180, 51-60.

Evans, C., Margison, F., & Barkham, M. (1998). The contribution of reliable and clinically significant change methods to evidence-based mental health. *Evidence Based Mental Health*, *1*(3), 70-72.

Goulding, R. & Goulding, M. (1976). Injunction, decision and redecision. *Transactional Analysis Journal, 6*, 41-8.

Iwakabe, S., & Gazzola, N. (2009). From single-case studies to practice-based knowledge: Aggregating and synthesizing case studies. *Psychotherapy Research, 19*(4-5), 601-611. DOI:10.1080/10503300802688494

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*(1), 12-19. DOI:10.1037/0022-006X.59.1.12

Kerr, C. (2013). TA Treatment of Emetophobia – A Systematic Case Study – 'Peter'. *International Journal of Transactional Analysis Research*, *4*:2, 16-26.

Kahler, T. (1975). Drivers: the key to the process of scripts. *Transactional Analysis Journal, 5,* 280-284.

Karpman, S. (1968). Fairy tales and script drama analysis. *Transactional Analysis Bulletin, 7*(26), 39-43.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Arch Gen Psychiatry*.62: 593–602. pmid:15939837 DOI: 10.1001/archpsyc.62.6.593

Klein, D. N. (2010). Chronic depression: diagnosis and classification. *Curr Dir Psychol Sci.*;19: 96–100.

Kroenke, K., Spitzer, R. L., Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. ;16(9):606-13

Kroenke, K., Spitzer, R. L., Williams, J. B., et al. (2007). Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Ann Intern Med*.;146(5):317-25.

Llewelyn, S. (1988). Psychological therapy as viewed by clients and therapists. *British Journal of Clinical Psychology*, *27*, 223-238. DOI:10.1111/j.2044-8260.1988.tb00779.x

MacLeod, R., Elliott, R., & Rodger, B (2012) Process-experiential/emotion-focused therapy for social anxiety: A hermeneutic single-case efficacy design study, *Psychotherapy Research, 22*:1, 67-81, DOI: 10.1080/10503307.2011.626805

McLeod, J. (2010). *Case study research in counselling and psychotherapy*. London: Sage Publications.

McLeod, J. (2013a). Process and outcome in pluralistic transactional analysis counselling for long-term health conditions: A case series. *Counselling and Psychotherapy Research, 13*:1, 32-43.

McLeod, J. (2013b). Transactional Analysis Psychotherapy with a woman suffering from Multiple Sclerosis. A Systematic Case Study. *Transactional Analysis Journal, 43*:3, 212-223.

NICE. (2009). *Depression: the Treatment and Management of Depression in Adults (Update).* NICE clinical guideline 90. Available at www.nice.org.uk/CG90 Accessed 10 January 2017

Norcross, J. C. (2002). *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients.* Oxford: Oxford University Press.

Ohlsson, T. (2010). Scientific Evidence Base for Transactional Analysis in the Year 2010. *International Journal of Transactional Analysis Research 1*:1 4-29

Richards, D. A. and Borglin, G. (2011). Implementation of psychological therapies for anxiety and depression in routine practice: two year prospective cohort study. *Journal of Affective Disorders, 133,* 51–60.

Shedler, J. & Westen, D. (1999). The Shedler-Westen Assessment Procedure (SWAP): Making personality diagnosis clinically meaningful. *Journal of Personality Assessment*, 89, 41-55.

Spitzer, R. L., Kroenke, K., & Williams, J. B. (1999). Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Journal of the American Medical Association. Nov 10; 282:18, 1737–44. PMID 10568646

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, *166*:10, 1092-1097. DOI:10.1001/archinte.166.10.1092.

Steiner, C (1966) *Scripts People Live,* New York: Grove Press

Stiles, W. B., Hill, C. E., & Elliott, R. (2015). Looking both ways. *Psychotherapy Research*, *25*(3), 282-293.

Wall, J. M, Kwee J. L, Hu, M. & McDonald, M. J. (2016). Enhancing the hermeneutic single-case efficacy design: Bridging the research–practice gap. *Psychotherapy Research, DOI: 10.1080/10503307.2015.1136441*

Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin, 130*:4, 631-663.

Widdowson, M. (2012a). TA Treatment of Depression - A Hermeneutic Single-Case Efficacy Design Study - 'Peter'. *International Journal of Transactional Analysis Research*, *3*:1, 1-11.

Widdowson, M. (2012b). TA Treatment of Depression - A Hermeneutic Single-Case Efficacy Design Study - 'Denise'. *International Journal of Transactional Analysis Research, 3*:2, 3-14.

Widdowson, M. (2012c). TA Treatment of Depression - A Hermeneutic Single-Case Efficacy Design Study - 'Tom'. *International Journal of Transactional Analysis Research, 3*:2, 15-27.

Widdowson, M. (2013). TA Treatment of Depression - A Hermeneutic Single-Case Efficacy Design Study - 'Linda'- a mixed outcome case. *International Journal of Transactional Analysis Research*, *4*:2, 3-15.

Widdowson, M. (2014). Transactional Analysis Psychotherapy for a Case of Mixed Anxiety & Depression: A Pragmatic Adjudicated Case Study – 'Alastair'. *International Journal of Transactional Analysis Research*, *5*:2, 66-76.

Widdowson, M. (2016). *Transactional Analysis for depression: A step-by-step treatment manual.* Abingdon: Routledge.