



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

© 2016 Enrico Benelli, Francesco Scottà, Serena Barreca, Arianna Palmieri, Vincenzo Calvo, Guido de Rénoche, Stefano Colussi, Marco Sambin & Mark Widdowson

Abstract

This study is the second of a series of three, and represents an Italian replication of a previous UK -based case series (Widdowson 2012a, 2012b, 2012c, 2013) that investigated the effectiveness of a recently manualised transactional analysis treatment for depression with British clients, using Hermeneutic Single-Case Efficacy Design (HSCED). The various stages of HSCED as a systematic case study research method are described, as a quasi-judicial method to sift case evidence in which researchers construct opposing arguments around multiple sources of quantitative and qualitative evidence and judges evaluate these to conclude whether the client changed substantially over the course of therapy, and whether the outcome was attributable to the therapy. The therapist in this case was a white Italian man in the third year of training to become a psychotherapist, and the client, Penelope, was a 45-year old white Italian woman with mild depression and anxiety. The conclusion of the judges was that this was a mixed-outcome case: the client improved some aspects of her problems, without obtaining a complete and stable remission. Interestingly, this case presents a minimal correlation between empirical and proxy-rated indexes of depression and anxiety and answers to self reported questionnaires, raising the question of validity of self report measures with specific typology of client.

Key words

Systematic Case Study Research; Hermeneutic Single-Case Efficacy Design; Transactional Analysis Psychotherapy; Depression; self report validity

Editor's Note

Those who read all three papers in this issue will see that some parts of the introduction, ethical considerations, method, and similar material, is repeated here for completeness of each paper.

Introduction

This article is the second of a series of three and represents an Italian systematic replication of previous UK-based findings (Widdowson 2012a, 2012b, 2012c, 2013) investigating the effectiveness of transactional analysis (TA) treatment of depression. This case was run under the auspices of the project 'Toward a transactional analysis psychotherapy recognised as empirically supported treatment: an Italian replication series design', funded by the European Association of Transactional Analysis (EATA). This present case study analyses process and outcome of the brief treatment of 'Penelope', a 45-year old Italian woman presenting with depression. The psychotherapy was conducted according to a recently manualized TA treatment of depression (Widdowson, 2015).

The general aim of this study was to investigate the effectiveness of short-term TA treatment of depression in a naturalistic setting. This article is interesting also because: the treatment is carried out by a young (28 year-old) psychologist in the third year of specialisation in psychotherapy; there is no correlation between the client's self-reported measures of depression and anxiety and the therapist's proxy measures and judgement; and the outcome is mixed, with quantitative and qualitative data presenting conflicting pictures of change. Therefore, this study addresses multiple important issues that practitioner-researchers come across in their routine practice. These include issues related to the quality of therapy delivered by trainee psychotherapists within public services, the validity and reliability of self-report measures, and how to interpret data that presents an ambiguous picture of change, where the first examination of the results suggests no clear conclusion regarding outcome. This is perhaps the sort of situation where HSCED shows its particular strength, in the detailed

analysis and cross-examination of qualitative and quantitative data and developing arguments that account for complex pictures of change which are then subjected to further analysis and evaluation by a process of adjudication to draw conclusions regarding the outcome of the case.

TA is a widely practiced form of psychotherapy, but it is still under-recognised within the worldwide scientific community of psychotherapy. Although its clinical efficacy is experienced in the consulting room by thousands of Transactional Analysts every day, systematic empirical evidence regarding the effectiveness of TA with specific presenting problems has been scant and of poor quality until recent years (Khalil, Callaghan & James, 2007). Ohlsson (2010) provided a valuable reference list of TA research studies. In order to define TA psychotherapy as an efficacious Empirically Supported Treatment (EST), its efficacy must have been established in at least one Randomized Clinical Trial (RCT) replicated by two independent research groups, or alternatively in at least three Single Case Efficacy Design studies (SCED), replicated by at least three independent research groups (Chambless & Hollon, 1998), with each group conducting a case series of a minimum of three cases. Recently, a wide community of researchers have proposed that treatment efficacy in psychotherapy is a complex issue which cannot be adequately evaluated with the experimental approach of either RCT design (Norcross 2002; Westen, Novotny & Thompson-Brenner, 2004) nor SCED alone (McLeod, 2010). Systematic case study research has been proposed as a viable alternative to RCT and SCED (Iwakabe & Gazzola, 2009), and 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 in psychotherapy (McLeod, 2010). HSCED is a systematic case study research method which examines individual cases and can be used to: (a) evaluate whether change has occurred; (b) examine evidence causally linking client change to the therapy; (c) evaluate alternative explanations for client change; and (d) identify the specific processes that appear to have been responsible for change (Elliott, et al., 2009).

Recently, a systematic review of all published HSCED studies found within English language peer-reviewed journals highlighted methodological issues related to different levels of stringency, offering solid alternatives according to the availability of resources for research (Benelli, De Carlo, Biffi & McLeod, 2015).

Systematic case study research has already been applied to investigate TA effectiveness with people with long term health conditions (McLeod, 2013a; 2013b) and HSCED methodology have been already successfully applied to TA and widely described in this Journal by Widdowson (2012a). Recently, several HSCEDs supporting TA treatment for depression (Widdowson, 2012a, 2012b, 2012c, 2013) 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 the effectiveness of TA for emetophobia (Kerr, 2013). The case series by Widdowson has shown that TA can be an effective therapy for depression when delivered in routine clinical practice, in private practice settings, with clients who actively sought out TA therapy and with white British therapist and client dyads.

Ethical Considerations

The research protocol follows the indications of the ethical code for Research in Psychotherapy of the Italian Association of Psychology and the American Psychological Association guidelines on rights and confidentiality of research participants. Before entering the treatment, the client received an information pack, including the detailed description of the research protocol, and gave her informed consent and written permission to use anonymised, disguised transcripts of segments of sessions or interviews within scientific articles and/or be presented at conferences. The client was informed that she would still have been entitled to attend therapy even if she decided against participating in the research and that she had the right to withdraw from the study at any time.

All aspects of the case material have been disguised, so that neither the client nor third parties are identifiable. All changes to anonymise the case have been made in such a way that would not lead the reader to draw false conclusions related to the described phenomena. Finally, this article, in Italian, was presented to the client following which she gave written consent for its publication.

Method

Inclusion and exclusion criteria

Participating psychotherapists were invited to include in the study the first new client with a diagnosis of depression who had agreed to participate in the research. The exclusion criteria for participation were: client in other current psychotherapy, active psychosis, domestic violence, bipolar disorder, antidepressant medication, and currently active alcohol or drug abuse.

Client

Penelope is 45 years old and lives with her 12-year old son, whom she adopted when he was 20 months old. She divorced approximately one year prior to the start of therapy. She consulted a neuropsychiatry clinic regarding the behaviour of her son, who has been diagnosed with Attention-Deficit/Hyperactivity Disorder (ADHD) and Oppositional Defiant Disorder (ODD). The psychologist of the service, noting that Penelope was tearful and clearly in palpable pain, suggested that she seek psychotherapy. Because of economic problems, she was referred to a psychologist specialising in psychotherapy, paying a small donation per session. At the first appointment, she spoke anxiously and quickly of many problems, crying often, especially when discussing

her relationship with her son, whom she described as overwhelming, insistent and demanding. She reported that he would often explode in uncontrolled anger when she tried to set boundaries for him. Given her son's reaction, she described how she starts to feel helpless and would like to get away, and then she feels incapable of managing the situation, and then begins a process of angry and guilty self-reproach connected to her feeling of not being able to manage. She described her situation as being unfortunate, and that no-one helps her, that she must carry out everything alone, and that she feels limited in every area of her life, stuck, frustrated, and unable to change things.

Penelope presented with a smart appearance and friendly manner. She described herself as generous, and that she does a lot for others, is able to give emotional support to others, and is surrounded by friends. She reported that in her friendships she feels confident and has fun. She described herself as a strong person who is able to find beauty in every aspect of her life, even when things go wrong. She tended to smile and cheerfully described painful or shocking events where she suffers severe abuse. Penelope reported that her work situation is good and that she has a good relationship with all of her colleagues.

She considers herself without any problems on a personal level, but overwhelmed by the external situations that she believes are unchangeable and therefore must be accepted. She appeared to be interested in participating in the therapy not to get personal change, but to help her child with his illness.

She described the recent divorce from her husband (about a year) after a long period of crisis, which began with the arrival of the child ten years earlier, which resulted in her ex-husband feeling that he was being overlooked. The ex-husband is described as physically violent, aggressive, demanding, both at work and with family members and relatives, unreliable in the role of parent and unable to care for the child in a responsible way (e.g., he does not pick up the son from school, did not accompany him to therapy or get involved in his sports activities). After the separation, he did not contribute economically, and indeed would repeatedly approach Penelope asking for money. Penelope reported that she always gave him money when asked, despite her own personal financial difficulties. She described his behaviour as intrusive, for example by calling her dozens of times a day. She described herself as resigned to suffer his behaviour and avoids telling him anything in order to prevent conflict.

Penelope briefly described that her father was an explosive and impulsive man who criticised others who expressed emotion. She described her mother as a person who becomes distressed every time she needs to make a decision.

Therapist

The therapist, a 28 year-old Italian man, was a psychologist at the third year of the four-year post degree training in psychotherapy. He was supervised at the end of each session by a Certified Transactional Analyst (Psychotherapy) (CTA P) with five years of experience and every four sessions by a Teaching & Supervising Transactional Analysts (Psychotherapy) (TSTA P) with more than 20 years of experience.

Intake sessions

Penelope attended two pre-treatment sessions which were focused on conducting a diagnostic interview evaluation according to DSM 5 criteria (American Psychiatric Association, 2013), developing a case formulation, creating a definition of the problems she was seeking help for in therapy, and for collection of self report outcome measure data relating to depression, anxiety and general psychological distress. The therapist also explained the research protocol and obtained informed consent from Penelope for her participation in this research. Penelope completed self-rated measures to assess her general suffering, depression and anxiety (see measures section). The scores of these measures were all within the 'healthy' range and did not indicate any clinical level of distress. However, the objective clinical examination revealed a depressive symptomatology. The therapist completed a proxy-rated measure to evaluate depression, generating a score which indicated moderate depression. Considering that Penelope tends to describe herself as a strong woman, who is used to doing everything alone without asking for help, it is possible that self-report measures did not reflect the true clinical situation, but were a reflection of her tendency to minimise her suffering. Due to the results of the diagnostic interview and the proxy-rated measure, the research team felt that including this case in the research project was justified.

DSM 5 Diagnosis

The therapist, during the diagnostic interview, on the basis of the objective examination and their clinical judgement, determined that Penelope met DSM 5 diagnostic criteria of Major Depressive Disorder. The therapist observed feelings of despair, hopeless and tearful (criterion 1) and psychomotor agitation (5), and Penelope described a diminished interest in almost all activities (2), increased appetite (3), fatigue and loss of energy nearly every day (6), feeling of worthlessness and inappropriate guilt (7) and diminished ability to make decisions (8).

Knowing the level of an individual's personality functioning and pathological traits provides the therapist with fundamental information for treatment planning. Therefore, a diagnosis of personality was also conducted, using the alternative dimensional model developed for DSM 5 Section III. This diagnosis allows:

1) assessment of the level of impairment in personality functioning and 2) an evaluation of personality traits. A moderate level of impairment in personality functioning is required for the diagnosis of a personality disorder, in at least two of the following areas: Identity, Self-direction, Empathy and Intimacy. The client showed some impairment in these areas, which resembled the prototypical description of the moderate level, leading to a diagnosis of moderate level of impairment in personality functioning. According to the new dimensional model of personality the client has a personality characterised by the Domains of Negative Affectivity (vs. Emotional Stability), more finely defined by the following Facets: Emotional lability, Anxiousness, Submissiveness, Depressivity, Distractibility. Both the level of personality functioning and personality traits were considered in drawing up the treatment plan.

TA Diagnosis and Case formulation

Penelope's depression was conceptualised as a consequence of painful internalisation of distress during her childhood that led to a pervasive sense of not being considered and supported by the other, which in turn led to a sense of inadequacy, inferiority and guilt. Penelope demonstrated a stroke filter (Steiner, 1974) that limited her perception of positive recognition. This supported the dominance of her powerful internalised critical Parent (Berne, 1967) and a lack of positive recognition from the protective nurturing Parent (Berne, 1964). She presented with Please Me and Be Strong drivers (Kahler & Capers, 1974), the injunctions (Goulding & Goulding, 1976) Don't be you, Don't be important, Don't feel and Don't express emotion (in particular anger and sadness). There was also some evidence of a Don't succeed injunction. Penelope's Racket System (Erskine and Zalzman, 1979) highlighted the following: Beliefs about self: "I am inadequate and inferior, weak and helpless"; Beliefs about others: "are better than me and will not help me"; Beliefs about life: "one must accept it as it is"; with repressed emotions of anger and sadness; Observable racket behaviours: 'pleasing' others, avoidance of loneliness; Reported internal experiences: some somatic symptoms, loss of energy; Reinforcing memory: archaic episodes of being criticized when showing suffering. Interpersonally, Penelope tended to present from a Victim role within the drama triangle (Karpman, 1968) and would counteract her sense of abandonment by then assuming a Nurturing Parent position and Rescuer role, taking care of others in order to keep them close.

Contract

Penelope entered therapy with the sole purpose of helping her child with his suffering. During contract negotiations Penelope focused on external issues (her son's illness, the behavior of her husband) and was less involved in exploring her own thoughts, emotions, needs, desires. After some negotiation, the final contract was that she would learn how to protect herself, to express

her own needs, and set limits and boundaries with others who were acting abusively towards her (primarily, her ex-husband).

Treatment

The therapy followed the manualised treatment of depression as described by Widdowson (2015), and tailored according to the level of personality functioning and personality traits in line with guidance from the treatment manual. To deal with a moderate level of impairment requires a permanent focus on the therapeutic alliance, creating a climate of Permission (Crossman, 1966) in contrast to the client's received injunction(s), providing a safe setting where she could feel and express repressed emotions, to explore her needs, to be herself, to explore options (Karpman, 1971), to change her behaviours and protect herself from others. The personality trait of Submissiveness requires a focus firstly on Protection (Crossman, 1966) in order to ensure that the client is not inadvertently exposed to violence or abuse, and is then followed by a redecision (Goulding & Goulding, 1976) about expressing her needs rather than avoiding conflict.

Analysis Team

The HSCED main investigator and first author of this paper is a Certified Transactional Analyst with 5 years of post-specialisation experience, with a strong allegiance to TA. Following the indication of Bohart (2000), the analysis was carried out by a team of 8 'reasonable persons', not yet overly committed to any theoretical approach or professional role. They were postgraduate students who were taught the principles of hermeneutic analysis in a course on case study research at the University of Padua, by Professor John McLeod. The students were split into two groups, the affirmative case and the sceptic case, with each group independently preparing their responses to the case. The main investigator supervised the briefs and rebuttals from both analysis teams.

Judges

The judges were two researchers in psychotherapy at the University of Padua and co-authors of this paper: Vincenzo Calvo, a psychologist and counsellor with expertise in attachment theory, and Arianna Palmieri, a neuropsychologist and psychotherapist with a training in dynamic psychotherapy. Both judges had some basic knowledge of TA but had not engaged in any official TA training.

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 in the final preparation of this article.

Quantitative Outcome Measures

Three standardized self-report outcome measures were selected to measure target symptoms: the Patient Health Questionnaire 9-item measure for depression (PHQ-9) (Spitzer, Kroenke & Williams, 1999), the Generalized Anxiety Disorder 7-item measure (GAD-7) (Spitzer, Kroenke, Williams, & Löwe, 2006) for anxiety and the Clinical Outcome for Routine Evaluation - Outcome Measure (CORE-OM; Phase 0, session 1, 8, 16 and follow-ups) (Evans, Connell, Barkham, Margison, Mellor-Clark, McGrath, & Audin, 2002) and short form (CORE-18; Pair session: short form A; despair sessions: short form B) (Evans, Mellor-Clark, Margison, Barkham, Audin, Connell & McGrath, 2000) for assessment of global functioning and distress. These measures were evaluated according to clinical significance and Reliable Change Index (RCI) (Jacobson & Truax, 1991). Clinical significance indicates that the client has moved from a clinical to a non clinical range score. RCI means that the change is reliable and not due to measurement errors. See Table 1 for RCI values for each measure.

All these measures were administered prior to 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 simplified 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. Furthermore the therapist compiled after every session the Hamilton Depression Rating Scale (HAM-D; Hamilton, 1960). All of these measures were administered also during the assessment phase to obtain a three-point baseline, and during the three follow-ups.

Qualitative Outcome Measurement

The client was interviewed one month after the conclusion of the therapy using the Change Interview protocol (CI) (Elliott, Slatick & Urman, 2001). The CI is a semi-structured qualitative change measure which asks clients how they feel they have changed during the therapy and since the start of the therapy, 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 therapy 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 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 and the two supervisors 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). Both supervisors compiled the adherence to treatment form and they indicated that the treatment was consistent with the TA theory to an adequate level of competence, with a slight improvement needed.

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 quantitative and qualitative measure. Change should be Clinically Significant (scores fall into 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 not present within the case, 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 index (Jacobson & Truax, 1991) 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 due to new relationship, work, financial conditions;
7. Psychobiological causes which verify whether change is due to medication, herbal remedies, recovery from medical illness;
8. Reactive effects of research, analysing the effect of change due to participating in research, such as generosity or good will towards the therapist.

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.

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.

Adjudication Procedure

Each 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 position.

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 factors.

Results

In earlier published HSCED 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, we adopted here the solution of providing a summary of the main points, as proposed in MacLeod and Elliott (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

Penelope's quantitative outcome data are presented in Table 1. Penelope's initial scores were well below the 'caseness' cut-off range for inclusion in this study: her CORE was 6.8, PHQ-9 was 3 and GAD-7 was 4, all indicating non-clinical range or healthy condition. The PQ score was 4.6, indicating a moderately to considerably bothering level of the problems for which Penelope was seeking help. The proxy evaluation of depressive symptomatology made by the therapist through HAM-D had a score of 16, indicating moderate depression, above the caseness cut-off.

During the therapy sessions, the self rated scores of CORE, PHQ-9 and GAD-7 and PQ appear stable, with slight fluctuation without Reliable Change, whereas the proxy-rated score of HAM-D showed a decrease from 14 to 9 (indicating lower limit of mild depression) by Session 8, and improvement that continued until the end of the therapy, where Penelope was rated as having a score of 2, which is well below the clinical cut-off and indicates recovery to normal.

At the first Follow Up, we observe a deterioration in all four self-rated measures, with CORE and GAD-7 slightly above the clinical cut off. Of these four measures, only GAD-7 shows a reliable deterioration according to the RCI. Also the therapist HAM-D score demonstrated a deterioration.

At the second Follow Up, we observe an improvement in all self rated measures: CORE and GAD-7 show a clinically significant and reliable change, and even the PQ for the first time shows a clinically significant change. The therapist HAM-D shows a slight, non-significant improvement.

Finally, at the third Follow Up, all self and proxy measures tend to deteriorate: PHQ-9 and GAD-7 show

reliable deterioration, with GAD-7 scores that are once again over the clinical cut-off threshold. The CORE and PQ scores also show a reliable deterioration. The therapist HAM-D shows a score of 12 (indicating the upper limit of mild depression).

The problems that the client identified at the beginning of therapy in her Personal Questionnaire are reported in Table 2. Figures 1, 2 and 3 show respectively the CORE, the PQ and the HAM-D weekly scores.

Qualitative Data

Penelope compiled the HAT form at the end of every session (Table 3), reporting only positive/helpful events within sessions, and with almost all sessions rated as 8 (very useful) with one rated 7 (useful). The first HAT form is missing.

Penelope participated in a Change Interview one month after the conclusion of the therapy, where she identified her main and significant changes that she felt were due to the therapy (Table 4). The first and the last reflect a behavioural change, whereas the others represent a change in self-perception. The researcher invited Penelope to describe her perception of the mechanisms of change and to what she attributed these changes. Penelope explained that she felt comfortable with the therapist and attributed her change to the ability of the therapist to support her disclosures (in the transcript, line P8), the ease with which she felt able to talk about her problems, and that the therapist provided a different perspective to that of her friends and was able to identify different view points (P26). She learned to think about how to put boundaries in place in order to protect herself and decide when other people were behaving abusively towards her (P21), to reduce the amount of her spare time that she had devoted to other people's needs (P28), and to be less passive and more assertive with others (P33). She felt supported by the therapist in expressing her emotions and thoughts, and in developing her ability to express her thoughts and feelings to other people, instead of her previous pattern of avoiding (P72). She felt that the therapist gave her a lot of space to express her feelings, without judgment or suggesting solutions, instead allowing her to find her own solutions (P83). On the other hand, Penelope suggested that her change may be due to the relationship with her new partner (P21) and that independently from the therapy, the new relationship gave her another kind of energy (P94). She also reiterated that most of her problems are due to other people (P33) and that her external problems are all still there (P94). Penelope did not identify any negative or unhelpful aspects of the therapy in her Change Interview. The only aspects that were considered negative by the client were the distance between the therapy practice and her home (about 20 km), the cost of the babysitter and the time required to travel to and attend sessions.

HSCED Analysis

Affirmative Case

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

The first line of evidence takes into account changes in stable/long-standing problems. As for quantitative data, in Table 1 we observe a mixed picture, with self-report outcome measure scores at the beginning scores within the 'healthy' range for functioning and distress (CORE), depression (PHQ-9) and anxiety (GAD-7), that are maintained with slight, but not reliable fluctuation until the end of the treatment phase (session 16). Despite this reported 'healthy' condition, the client refers to problems which bother her at the moderately to considerably level in her PQ. Indeed, the perception of the clinician assessing Penelope was of a woman who was clearly experiencing distress. This is reflected in the therapist's initial evaluation of moderate depression (HAM-D), that from Session 13 demonstrated a clinically significant (and reliable) change. This mixed picture suggests that the scores might be biased by social- and self-presentation effects. This is a well-known phenomenon, reported in the scientific literature by psychotherapists and physicians, and applies to a group of people who present low scores on self-report measures, but reveal evidence of significant suffering when assessed in other ways. For this reason, it appears that in this case the clinical judgment of the therapist may be more reliable than Penelope's self-report.

Qualitative data supports this conclusion: Penelope in her Change Interview recognised that she can change things for herself, even if others do not change. This aspect represents a clear break with her initial statement, when she stated that all of her problems had an external cause or were determined by others, and that she could only accept the situations as they were.

We observe a deterioration at the first follow up in all self- and proxy-rated measures: CORE and GAD-7 passed into clinical range indicating the emergence of mild distress and anxiety, and the therapist HAM-D rating is barely beneath the clinical range. This appears a transient effect of the conclusion of the therapy, that occurred during a period of distress for Penelope, who described difficulties in her relationship with her new partner.

In fact, at the second follow up we observe an overall improvement, with self-report measures showing the best scores since the beginning of the therapy. The PQ shows a reliable change, and is around the clinical cut-off level. During the second follow up, Penelope described herself as more able to think for herself, to feel less guilty and as more aware of her needs.

	Clinical Cut-Off	Case Cut-Off	Reliable Change Index	Pre-Therapy	Session 8 (middle)	Session 16 (end)	1 month FU	3 months FU	6 months FU
CORE	10	15	5.1	6.8	7.4	6.5	10.8	3.8	9.1
PHQ-9	10	15	6	3	2	3	4	1	6
GAD-7	8	10	4	4	3	4	8	3	11
PQ	3	3.5	1	4.6	4.1	4.1	4.6	3.1 (*)	4.6
HAM-D	7	14	-	16	9	2 (+)	6 (+)	4 (+)	12

Table 1: Penelope's Quantitative Outcome Measure

Note. Values in **bold** are within clinical range; + indicates clinically significant change (CS). * indicates reliable change (RCI). CORE = Clinical Outcomes in Routine Evaluation-Outcome Measure (Evans et al., 2000). 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). HAM-D = Hamilton Depression Rating Scale (Hamilton, 1960). FU = follow-up.

	PQ Items	Duration	Pre-Therapy	Session 8 (middle)	Session 16 (end)	1 month FU	3 months FU	6 months FU
1	I feel helpless facing my son's crisis	3-5 years	4	4	4	5	3	4
2	I'm not able to help my son with homework	3-5 years	5	5	4	5	2	3
3	I'm not able to be obeyed by my oppositional son	6-10 years	5	5	5	6	4	5
4	I'm not able to feel peaceful	3-5 years	5	4	4	4	3	5
5	I feel alone managing my son's problems	3-5 years	7	5	6	6	6	7
6	I feel stressed	1-2 years	4	3	3	4	3	5
7	I'm not able to set limits to my ex-husband	3-5 years	3	5	5	4	3	4
8	I feel up to the eyeballs	3-5 years	4	2	2	3	1	4
	TOTAL		37	33	33	37	25	37
	MEAN		4,6	4,1	4,1	4,6	3,1	4,6

Table 2: Penelope's Personal Questionnaire items

Note: Values in **bold** are within clinical range. The rating is on a scale from 1 to 7 and indicate how much each problem has bothered the client during the previous week: 1 = not at all; 9 = completely. FU= follow-up

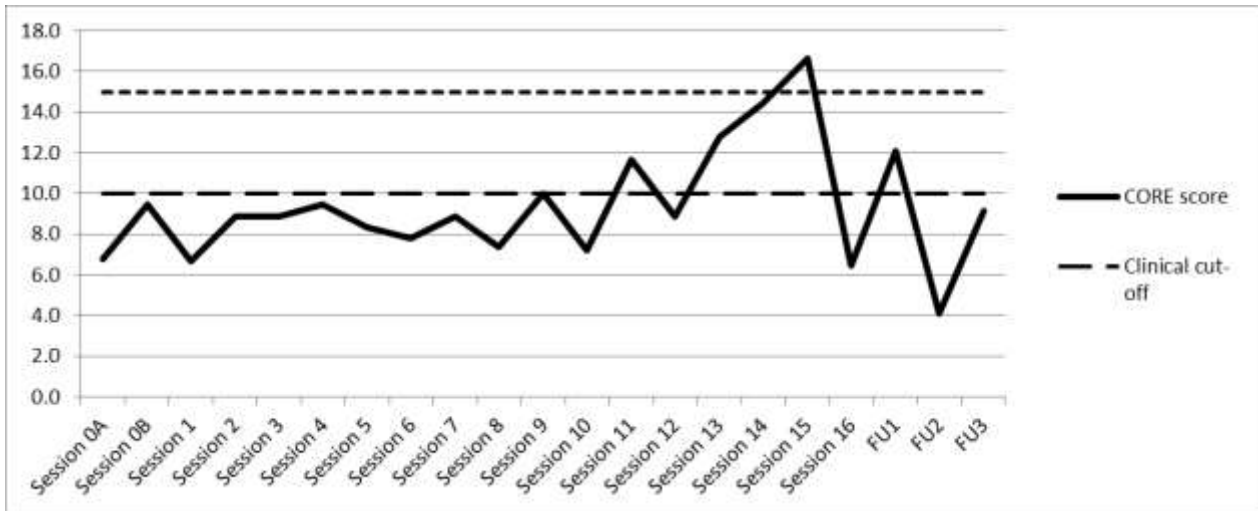


Figure 1: Penelope's weekly CORE score

Note. 0A and 0B = assessment sessions. CORE = Clinical Outcomes in Routine Evaluation-Outcome Measure (Evans et al., 2000). FU = follow-up.

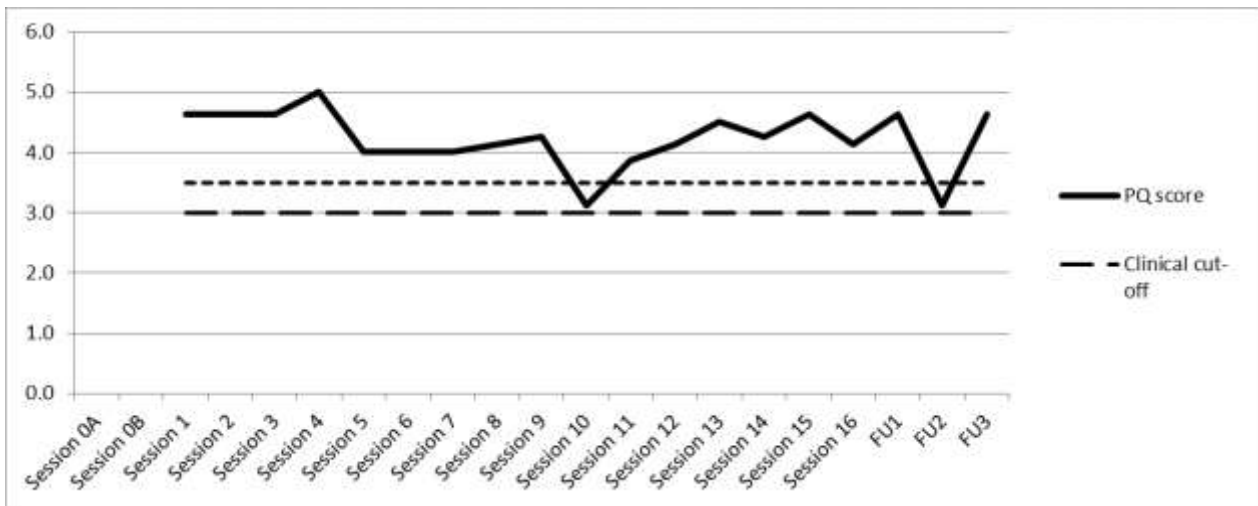


Figure 2: Penelope's weekly PQ score

Note. 0A and 0B = assessment sessions. PQ = Personal Questionnaire (Elliott, Shapiro, & Mack, 1999). FU = follow-up.

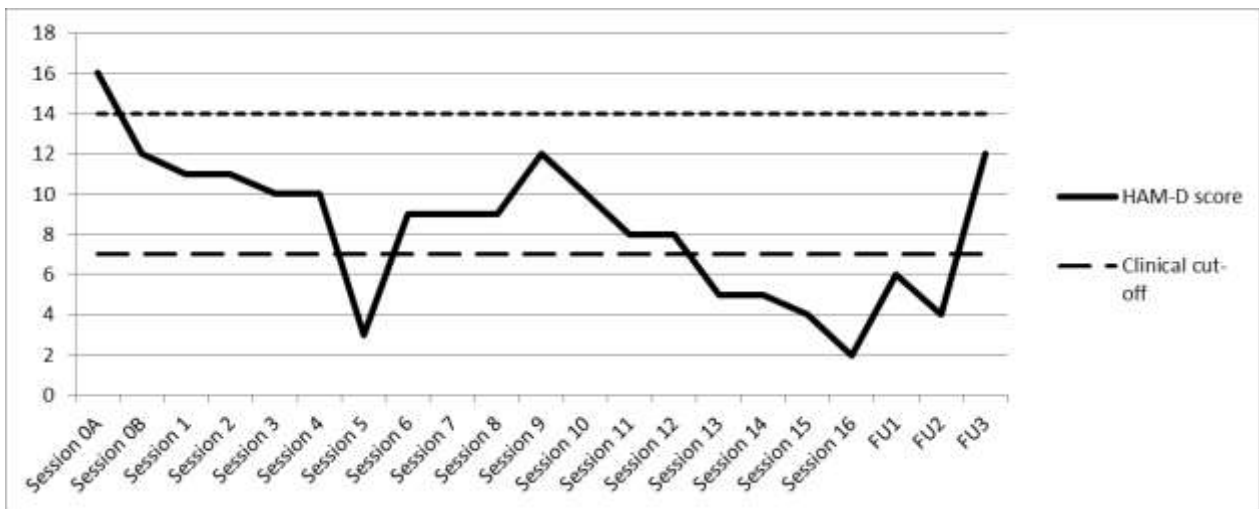


Figure 3: Therapist's weekly evaluation of Penelope depression - HAM-D score

Note. 0A and 0B = assessment sessions. HAM-D = Hamilton Depression Rating Scale (Hamilton, 1960). FU = follow-up.

Session	Rating	Events / What made them helpful
1	-	missing
2	7	express my feelings / The therapist understood my difficulty
3	8	express "anger" for the scant help I receive from those who could and should help me (school/church); My change regarding the management of my child's homework / let off steam
4	8	reiterate that I have to face everything alone and that in any case I have energy enough to try to do something for myself
5	8	telling the meeting with my new lover / express my positive feeling, hindered by my sense of guilt
6	8	talk about my experiences with greater serenity / to be very relaxed in expressing my emotions
7	8	express my emotions and my "conflict" about the relationship / to speak peacefully
8	8	express my awareness about undergoing change / it was useful that the operator pointed out to me my change
9	8	to show the conflict between reason and emotions / I am aware that I decided according to emotions, but I know that I have also the clarity of the reason
10	8	express my conviction of preserving what I have received of beautiful and good from my lover / important because make me feeling more strong and secure
11	8	establish my boundaries / useful to rediscover my inner strength
12	8	express my new situation with my lover / I called into question once again a thought that I believed clear and unchangeable
13	8	talk about the difficulties with the school of my son / express my anger and frustration about the behaviour of headmaster and teachers
14	8	deal with the end of the relationship with my lover / unburden myself and express my emotions of suffering
15	8	recognise that I have to give me time to understand how to cope with new experiences / do not judge myself and have a little patience
16	8	link the management of the crisis of my child with his fear for my health / useful because I understood what was behind the anger of my son

Table 3: Penelope's helpful aspect of therapy - HAT

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

CI items	How much change was expected ¹	How /likely change would have been without therapy ²	Importance of change ³
To create boundaries in order to protect myself	surprising (5)	quite unlikely (2)	very (4)
not to get discouraged by the behaviour of my ex-husband	quite expected (2)	quite unlikely (2)	very (4)
To self guard from others	surprising (5)	quite unlikely (2)	very (4)
To think about my self	neither (3)	unlikely (1)	very (4)
If others do not change, I have to	surprising (5)	quite unlikely (2)	very (4)

Table 4: Penelope's changes identified in Change Interview – CI (Elliott et al. 2001)

¹The rating is on a scale from 1 to 5; 1= expected, 3= neither, 5= surprising. ² The rating is on a scale from 1 to 5; 1=unlikely, 3=neither, 5=likely. ³ The rating is on a scale from 1 to 5; 1=slightly, 3 = moderately, 5=extremely.

Despite this, at the third follow up all measures once again worsen. The PHQ-9 and the GAD-7 show the worst score ever, the PHQ-9 does not reach the clinical cut off, whereas the GAD-7 reached a score of 11, indicating moderate anxiety. The CORE touches on the clinical cut-off, with a score higher than the pre-therapy level, and the PQ returns to the pre-therapy score. Also the therapist rated the HAM-D with a score of 12 which indicates a relapse into mild depression. Deterioration at the third follow up appears to reflect adverse circumstances. Penelope described in particular experiencing some problems at work, with uncertainty around her future, together with many difficulties with her son who had just started middle school, and her troubled relationship. These problems all challenged her previous tendency of feeling overwhelmed by events that she feels she has no control over and that are due to external conditions.

Retrospective attribution - Penelope in the Change Interview (Table 4) describes five changes, all rated as very important. Penelope also reports that all changes would be quite unlikely or unlikely without therapy. She felt surprised that she had been able to put limits on others in order to protect herself (item 1), and realising that she is able to make personal changes, even if others do not change (5). These descriptions show an overall change in representations of self and others, self-esteem and relational patterns. She recognizes the role of the therapist when she states "I already spoke about my problems with my friends, but with the therapist I approached them in another way... He encouraged me to notice other aspects" (P 26). She affirms "the therapist got me thinking about how to set limits... I mean, a kind of protection for me." (P21). She also recognises that now "I have learned that if things that are wrong are not

changed by others, I must change them: just the question of setting limits with others. I mean... if they do not understand what... I want, mmh... it is logic I have to change".

Association between outcome and process (outcome to process mapping) - Penelope in her HAT forms (Table 3) reported several within session events that she rated 8, very useful (but the first, rated 7, moderately useful). In the first seven sessions, she describes that she felt understood and able to express her feelings, which is related to the therapist notes that report a focus on alliance and leaving room to express her feelings, providing permissions for Penelope to be herself and express her emotions. This appears to be tied in the Change Interview to the unexpected change of "thinking about my self" (item 4), instead of over-adaptation to others (as reported in the change interview, P 28). In the following sessions, the therapist's focus were on Penelope's needs, internal conflicts and how to express her feeling and needs to others. This also reflects her description of relationships gathered in the HAT forms about her former husband (11), her son (16), her new partner (10, 14) and the teachers of her son (13). These interpersonal change are reflected in the Change Interview in the first three items. It appears particularly important for Penelope and the relationship with her son that she started to understand the link between the aggressive behaviours of her son and his underlying emotion of fear (16).

Event-shift sequences – It appears that the change observed in Penelope is not tied to a pivotal moment in the therapy, such as a single session or technique, but appears related to the general relational climate experienced during the treatment.

Within therapy process-outcome Correlation - As for the fifth source of evidence, no correlation between within-therapy processes measure by the adherence form and quantitative outcome measures has been found, suggesting global rather than temporary change.

Conclusions - Although quantitative self reported data on depression, anxiety and global distress appear to be unreliable, a proxy-rated measure of depression and qualitative data from the Change Interview support the claim that Penelope experienced a partial good outcome during the therapy, that appeared to have grown at the second Follow Up but was not maintained at the third Follow Up. The deteriorated scores on both proxy- and self-reported measures observed at the third Follow Up appear to be due to several external factors, and suggest that the client developed a major awareness about herself and her suffering in line with the therapist's perception, whereas at the beginning of the therapy her suffering was bodily expressed but not reported on questionnaires. The change appears linked to the treatment, particularly focused on empathic listening, permission to feel and express emotions and needs, analysis of the critical internal dialogue and the exploration of more self-protective options within relationships with others.

Sceptic Case

In Table 1 it appears that Penelope does not go above the caseness cut-off for any self reported measures, suggesting that she did not meet the inclusion criteria of the study. Even accepting that some social or personal factors may have biased the initial scores, we observe a negative change between the scores of the pre-therapy and the third Follow Up, suggesting a possible iatrogenic effect of the therapy. The therapist's rating of depression (Figure 3) shows a three-point decrease from baseline between the beginning of the assessment phase and the first Session included, suggesting that there is not a stable baseline from which the following results can be adequately compared. Furthermore, to rely solely on therapist-rated measures introduces a risk of experiment-wise errors, since the validity of his score may be biased by social and personal factors.

The apparently positive results observed at the second Follow Up (scheduled in mid June) might be tied to the end of the school year for Penelope's son and the start of the summer holidays; both external factors may have reduced Penelope's initial perception of being overwhelmed (as reported in the Personal Questionnaire items) and her distress. This would also explain why at the third Follow Up (in September) we observed deterioration which Penelope attributed to difficulties with her son's school and her work environment.

Qualitative data suggests a more positive picture, but from the fifth Session Penelope describes being in love with a new partner, making it difficult to differentiate between the effect of this event and the effects of the therapy on her self esteem and behavioural change with

her ex-husband and other persons. Furthermore, in her Change Interview Penelope affirms "surely the therapy helped me, but an event happened to me just at the beginning of the therapy... in the same period at a human level... I knew a person that... made me change" (P 21). Also, she affirms "I cannot quantify how much my problems are changed... I mean, my entire problems are still there... I mean, not that there are none, because unfortunately they depend on the people around me" (P 94), suggesting that the change, if any, is due to external factors and that her attribution style is not changed.

Conclusions - The sceptic case concludes that this therapy had a poor outcome. Quantitative self-reported measures show that the change is trivial and potentially negative, and quantitative proxy-rated measures of depression are biased and flawed due to the lack of a stable baseline. The improvement observed at the second follow up appears to be due to the start of the summer holiday period. Qualitative data supports the view that any positive change reported is due to external factors, such as the new relationship, and the summer vacations.

Affirmative Rebuttals

Although quantitative self-report measures did not justify the inclusion of Penelope in this research, inclusion criteria should not take precedence over clinical judgment. Self-report measures are biased by the same social and personal factors as proxy rated measures. The affirmative team believe that this argument raises a question about the usefulness of inclusion criteria, if clients like Penelope are considered healthy and therefore excluded from a study. An early decrease in depressive symptoms is observed also in clients who are on a waiting list, and this is associated to an effect associated with hope. It is not surprising that Penelope shows an early improvement of her symptoms, and this may support the accuracy of the therapist rating. The HAM-D score shows a substantial improvement in the fifth session, which corresponds with the beginning of the relationship with her new partner, again supporting the accuracy of the therapist observations.

The affirmative team concede that it is possible that the summer holidays had a beneficial effect on Penelope's distress, as well as the start of her new relationship, but also winter holidays occurred between the eighth and ninth Sessions, and Easter holidays too, immediately after the first Follow Up, neither of which appeared to have any positive effects on the self reported measures. Instead, it appears from transcriptions that during the therapy, Penelope got in touch with an increasing level of distress: growing tensions with her ex-husband, uncertainty about a possible redundancy, and even the new relationship which caused some pain. Such distress was discussed with the therapist, and had no impact on quantitative measures. We suspect that the ending of therapy contributed to a worsening of symptoms noticed during the follow-up period, but believe that an increase in her awareness about her level of distress was reflected

in her self-rated scores. In this sense, it is possible that her deterioration represents a more accurate perception of her own distress. The situation at the third follow up is described as more overwhelming than ever, with a deteriorating situation at work, with her ex husband and with her son. She feels helpless, unable to negotiate, and does not expect any change. Thus, the effect of the therapy has been present but is not sustained after the conclusion.

In her change interview, Penelope reports contradictory affirmations that brought about her change "I think that meeting this new partner is important" (P55), and that change "probably sooner or later would have happened the same, but maybe not so quickly" (P49), together with opposite affirmations such as "I understood I can understand situations in different ways, not in a passive way... having determination, not accepting criticism" (P 68) and "I feel stronger, I mean... more decisive, I can stop people... I can use a sort of protection" (P92). The overall meaning of this should be taken from a whole consideration of the Change Interview, HAT forms and transcripts of the session, rather than from a single sentence. Further, the tendency not to recognize her success is in line with her personality trait and diagnosis.

Sceptic Rebuttal

The problems reported by the client at the third Follow Up are not so severe, since they are common problems that any person can experience. The changes appear tied to the presence of the therapist and the client has not internalized sufficient resources to maintain change and deal with stress. Therefore, the therapy was not effective.

Adjudication

Each judge examined the rich case and hermeneutic analysis and independently prepared their opinions and ratings of the case (Table 6). The judges have considered that the quantitative data show an improvement in symptoms from the therapist's point of view, which has been confirmed in supervisions, but they show a deterioration from the client's point of view.

Moreover, they observed an inconsistency between quantitative and qualitative data, therefore both judges believe that this could be a mixed outcome case.

Summary of opinions regarding how the judges would categorise this case

According to Judge A (Calvo), the case appears mixed (60% of certainty) to poor (40% of certainty) outcome. Quantitative data do not allow for claims of a good outcome, and also qualitative information appears inconsistent and unduly influenced by significant external events.

According to Judge B (Palmieri) the case appeared to be a mixed (80% of certainty) to good outcome (20% of certainty). Neither quantitative or qualitative self-report data supports a clear claim of a positive outcome, but nonetheless it is possible to observe within the transcriptions of the sessions several relational episodes that indicate a change in behaviours, relationships and self-esteem. For example, the client is able to express herself more clearly to the husband and to the teacher and headmaster of the son's school; she recognises that her needs are as important as others' and reports a deeper comprehension about the needs of her son. This represent a clear change, even if not stable or complete.

Summary of opinions regarding the extent to which the client had changed

Judge A believes that this case presents a limited outcome, a moderate change with a certainty of 80%. There is some evidence of change, but the changes observed are not stable and are not maintained over time. The change does not appear in self-report instruments completed by the client and, at the beginning of the therapy, presented with sub-clinical scores for anxiety, depression and global distress. In her Change Interview the client refers to some important changes of her life, such as to identifying her own needs and an increased ability to protect herself by putting boundaries in place with others. In any case, the change is not maintained when the situation at workplace of the client

	Judge A	Judge B	Mean
How would you categorize this case?	Mixed to poor outcome	Mixed to good outcome	Mixed outcome
How certain are you?	60%	80%	70%
To what extent did the client change over the course of therapy?	40% moderately	60% considerably	50% moderately to considerably
How certain are you?	80%	80%	80%
To what extent is this change due to therapy?	60% considerably	80% substantially	70% considerably to substantially
How certain are you?	60%	80%	70%

Table 6: Adjudication results

becomes stressful and there were problems at her son's school. Therefore, although there are changes, they do not appear to be stable.

Similarly, clinical evaluation of the depression made by the therapist showed a significant improvement at the end of the treatment and in the first two Follow Up meetings, but is not maintained at the third Follow Up, confirming the hypothesis that the results achieved from the treatment are limited and not stable.

Judge B affirms that there was a considerable change, with 80% of certainty. Such affirmation is based on the consideration that the client showed some behavioural change in long-standing daily life situations such as the relationship with the husband and the son.

Summary of opinions as to whether the changes were due to the therapy

According to Judge A, the quantitative data shows a picture of limited change which is not lasting, although analysis of the qualitative data suggests significant changes occurred during therapy. Although the client stated in her Change Interview that the change could potentially be related to her new relationship, and the sceptic case also attributes change to her new relationship, these changes have not remained stable. Furthermore, the argument of the affirmative case that the client has learned to put boundaries in place with her new partner appears convincing and clinically meaningful. It appears unlikely that the change in the client's self-perception could occur without therapy. In addition, from the affirmative case perspective and from the transcripts of the sessions, what emerges is a clear relationship between the work carried out by the therapist and the perception of the client reported in HAT.

Judge B states that it appears improbable that such changes may be due only to external factors, and in particular that Penelope has learned "to express feelings, especially anger," recognise that she overcompensates due to feeling guilty for not having done enough, and deal with her critical internal dialogue between conflicting needs.

Both judges affirm that the changes are not of great magnitude, but are nevertheless meaningful and significant for the client and have an impact in her daily life with her son, her ex-husband, her new partner, her environment and her dealings with her son's school.

Mediator factors

The client seems to have experienced considerable benefit from the atmosphere of non-judgmental listening established within the sessions. The client describes the therapist's attitude "without judgment...he left much room...left the way you express yourself ... he gave no suggestions ...he (allowed for) things to come out". In addition, the therapist has created an atmosphere in which he accepted all of Penelope's emotions, giving continuous permissions to feel and express emotions; particularly anger and sadness. Also the focus on the

analysis of her negative internal dialogue between critical Parent and Child ego states also appears to have been important. Finally, the systematic exploration of new options for expressing needs and enforcing boundaries with others was important in this case.

Moderator factors

The positive effects of the therapy may have been moderated by the client's supporting network of friends, which potentially represented a protective factor against depression. As for negative moderator factors, Judge A pointed out that this case is characterised by low initial client motivation, low willingness to engage in treatment and that the therapy approach had not been actively sought out or chosen by the client. All these elements are known to be associated with poor outcome (Norcross, 2002).

Discussion

The general aim of this paper was to investigate the effectiveness of short-term TA treatment of depression in a naturalistic setting. The judges agreed that this is a mixed outcome case and Penelope did change as a result of the therapy, but the change is not sustained. It is therefore likely that this case does not add support to the recognition of TA as an Empirically Supported Treatment.

However, this case underlines some aspect of research in psychotherapy on which it may be useful to reflect. First of all, this case was conducted by a young psychologist in training to become a psychotherapist. In Italy, most of the therapy provided within public services are carried out by therapists at this same level of expertise. This study suggests that young therapists provide treatment at an adequate level of competence, even if some improvement is needed.

This case also raises questions about the reliability of self-report measures. Shedler and colleagues (1993) identified a group of people defined with the term 'illusory mental health' who tend to record low scores on self-report measures, but present with clear and genuine distress. These people do not generally appear in controlled clinical trials because they would be excluded due to not meeting inclusion criteria. According to McLeod (2001), the impact of an effective therapy on such clients is to increase scores in measures related to their denied suffering. Although the use of self-report measures dominates the field of psychotherapy research, there are several different points of view that are worthy of consideration. Firstly, self-report measures may be considered insensitive to situational factors that influence behaviour. A second potential problem is of minimal correlation between answers to a questionnaire and an empirical index that assesses the same variable such as overt behaviour or overt impairment in functioning (Holzman and Kagan, 1995). Where there is a high correlation between the variable (e.g. anxiety or depression) and social desirability ratings, people are likely to be influenced by the social desirability of each item whilst answering questionnaires (Edwards, 1957).

Thirdly, this case raises the question of how to evaluate qualitative and quantitative data that are inconsistent with each other. It appears clear, for example, that Penelope experienced during her therapy a kind of relationship that changed her comprehension of both her emotions and those of her son. A change like this, clearly reported in the HAT and within the session transcripts, appears relevant to a psychotherapist, but is usually not investigated by research design.

Limitations

This study may be biased by the role of the first author who is also the supervisor of the therapist and a teacher of the members of the hermeneutic groups, and who collaborates professionally with both judges. Despite the critical reflexive attitude and the auditing of expert researchers (Dr Mark Widdowson and Professor John McLeod), bias may nevertheless have influenced the results and analysis in subtle ways.

The baseline consisted of only two measurement intervals whereas international standards require at least three measurement intervals to make claims of a stable baseline.

Conclusion

The present case does not appear to support the effectiveness of TA short treatment of depression. However, it does provide an example of the various challenging situations that a practitioner can come across when trying to do research in his/her routine practice. A deep examination of session notes or transcripts allows an expert clinician to establish the gravity of a client's distress, and the magnitude of a change in a way that self-report cannot estimate. Despite this, the empirical support of many psychotherapy models is greatly influenced by findings taken from self-report measures. We do not wish to discount the importance of such questionnaires and self-rated measures, but believe it is important to enrich and integrate such measures with proxy rating, clinical judgement and qualitative data.

More research is needed to support the growing body of evidence relating to the efficacy and effectiveness of TA psychotherapy for depression, and to enable recognition of TA as an empirically supported treatment.

Funding

This study was supported by a grant from the European Association of Transactional Analysis, as part of the project 'Toward a transactional analysis psychotherapy recognised as empirically supported treatment: an Italian replication series design', and by a grant of the Center of Dynamic Psychology - Padua, a transactional analysis oriented School of Specialization in psychotherapy.

Enrico Benelli PhD, Certified Transactional Analyst (Psychotherapy), Vice-President of the Center of Dynamic Psychology of Padua (Italy), Adjunct Professor, University of Padua can be contacted at: enrico.benelli@unipd.it

Francesco Scottà, Psychotherapist, Center of Dynamic Psychology, Padua

Serena Barreca, Teaching & Supervising Transactional Analyst (Psychotherapy)

Arianna Palmieri, Assistant Professor, University of Padua

Guido de Rénoche, Neuropsychiatrist, Hospital Agency, University of Padua

Vincenzo Calvo, Assistant Professor, University of Padua

Stefano Colussi, Psychologist, Center of Dynamic Psychology, Padua

Marco Sambin, Certified Transactional Analyst (Psychotherapy), Full Professor, University of Padua

Mark Widdowson PhD, Teaching & Supervising Transactional Analyst (Psychotherapy)

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Benelli, E., Biffi, D., De Carlo, A., & McLeod, J. (2015). Hermeneutic Single Case Efficacy Design: a systematic review of published research and current standards. *TPM: Testing, Psychometrics, Methodology in Applied Psychology*, 22(1).
- Berne, E. (1964). *Games people play*. New York: Grove Press.
- Bohart, A. C. (2000). *A qualitative 'adjudicational' model for assessing psychotherapy outcome*. Paper presented at the meeting of the Society for Psychotherapy Research, Chicago.
- Chambless, D.L. and 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(19), 152-154.
- Edwards A. L. (1957) *The Social Desirability Variable in Personality Assessment and Research*. New York: Dryden.
- Elliott, R. (2002). Hermeneutic single-case efficacy design. *Psychotherapy Research*, 12, 1-21.
- Elliott, R., Shapiro, D. A., & Mack, C. (1999). *Simplified Personal Questionnaire procedure manual*. Toledo, OH: University of Toledo.
- 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, 543-557.
- Elliott, R., Slatick, E., & Urman, M. (2001). Qualitative change process research on psychotherapy: Alternative strategies. *Psychologische Beiträge*, 43, 69-111.

- Erskine, R. G., & Zalcman, M. J. (1979). The racket system: A model for racket analysis. *Transactional Analysis Journal*, 9(1), 51-59.
- Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., & McGrath, G. (2000). CORE: Clinical outcomes in routine evaluation. *Journal of Mental Health*, 9, 247-255.
- Evans, C., Connell, J., Barkham, M., Margison, F., McGRATH, G. R. A. E. M. E., Mellor-Clark, J., & Audin, K. (2002). Towards a standardised brief outcome measure: psychometric properties and utility of the CORE—OM. *The British Journal of Psychiatry*, 180(1), 51-60.
- Goulding, R., & Goulding, M. (1976). Injunctions, decisions, and redecisions. *Transactional Analysis Journal*, 6(1), 41-48.
- Hamilton, M. (1960). A rating scale for depression. *Journal of neurology, neurosurgery, and psychiatry*, 23(1), 56.
- Holzman, P., & Kagan, J. (1995). Whither or wither personality research. In P. E. Shrout, & S. T. Fiske (Eds.), *Personality research, methods, and theory: A Festschrift honoring Donald W. Fiske* (pp. 3-11). Erlbaum Hillsdale, NJ: Psychology Press
- 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.
- Kahler, T., & Capers, H. (1974). The miniscript. *Transactional Analysis Journal*. 4:1 26-42
- Khalil, E., Callaghan, P., & James, N. (2007). *Transactional analysis: A scoping exercise for evidence of outcome*. Report prepared for the Berne Institute. The University of Nottingham, School of Nursing.
- Karpman, S. (1968). Fairy tales and script drama analysis. *Transactional Analysis Bulletin*, 7(26), 39-43.
- Karpman, S. (1971). Options. *Transactional Analysis Journal*, 1(1), 79-87.
- Kerr, C. (2013). TA Treatment of emetophobia — A systematic case study — “Peter.” *International Journal of Transactional Analysis Research*, 4, 16-26.
- 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.
- Llewelyn, S. P. (1988). Psychological therapy as viewed by clients and therapists. *British Journal of Clinical Psychology*, 27(3), 223-237.
- MacLeod, R., & Elliott, R. (2012). Emotion-focused therapy for social anxiety: A hermeneutic single case efficacy design study of a low-outcome case. *Counselling Psychology Review*, 27, 7-22
- McLeod, J (2001): An administratively created reality: Some problems with the use of self-report questionnaire measures of adjustment in counselling/psychotherapy outcome research, *Counselling and Psychotherapy Research*, 1:3, 215-226. <http://dx.doi.org/10.1080/14733140112331385100>
- McLeod, J. (2010). *Case study research in counseling and psychotherapy*. London, UK: Sage.
- 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.
- Norcross, J. C. (Ed) (2002). *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients*. Oxford: Oxford University Press
- Ohlsson, Thomas (2010) Scientific Evidence Base for Transactional Analysis in the Year 2010 *International Journal of Transactional Analysis Research* 1:1 4-29
- Shedler, J., Mayman, M., & Manis, M. (1993). The illusion of mental health. *American Psychologist*, 48(11), 1117.
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Patient Health Questionnaire Primary Care Study Group. (1999). Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *Jama*, 282(18), 1737-1744.
- 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, 1092-1097.
- Steiner, C. (1974). *Scripts people live: Transactional analysis of life scripts*. Grove Press.
- 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.
- Widdowson, M. (2012a). TA Treatment of depression — A hermeneutic single-case efficacy design study — “Peter.” *International Journal of Transactional Analysis Research*, 3, 1-11.
- Widdowson, M. (2012b). TA Treatment of depression — A hermeneutic single-case efficacy design study — ‘Denise’ *International Journal of Transactional Analysis Research*, 3, 3-14.
- Widdowson, M. (2012c). TA Treatment of depression — A hermeneutic single-case efficacy design study — ‘Tom’ *International Journal of Transactional Analysis Research*, 3, 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, 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. (2015). *Transactional Analysis for depression: A step-by-step treatment manual*. Hove: Routledge