



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

© 2017 Enrico Benelli, Maddalena Bergamaschi, Cristina Capoferri, Stefano Morena, Vincenzo Calvo, Stefania Mannarini, Arianna Palmieri, Mariavittoria Zanchetta & Mark Widdowson

### Abstract

*This study is the third 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). Major Depression and Subthreshold Depression are often in comorbidity with Anxiety disorders in childhood and adolescence and represent a risk factor for ongoing mental health problems in adulthood. The therapist was a white Italian woman with 15 years of clinical experience and the client, Deborah, was a 15-year old white Italian female adolescent who attended sixteen sessions of transactional analysis psychotherapy. The conclusion of the judges was that this was a good-outcome case: the depressive and anxious symptomatology clinically and reliably improved over the course of the therapy and these improvements were maintained throughout the duration of the follow-up intervals. Furthermore, the client reported significant change in her post-treatment interview and these changes were directly attributed to the therapy. In this case study, the transactional analysis manualised treatment for depression in adulthood has demonstrated its effectiveness also in treating depressive and anxiety symptoms in adolescence.*

### Key words

Systematic Case Study Research; Hermeneutic Single-Case Efficacy Design; Transactional Analysis Psychotherapy; Major Depressive Disorder; Anxiety Disorder; Adolescence.

### Introduction

This study is the third 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 et al., 2016a, 2016b, 2016c)

and was conducted under the auspices of the European Association for Transactional Analysis (EATA) and the University of Padua.

Major Depressive Disorder (MDD) affects all age groups, and is considered the fourth leading cause of disability in Europe and North America, calculated by Disability Adjusted Live Years (DALYs) (Murray, Vos, Lazano et al., 2012). Depression in childhood and adolescence has an estimated prevalence of 2,8% amongst children and 5,9% amongst adolescents (Costello, Erkanli & Angold, 2006). This figure rises to almost 25% of adolescents attending primary care settings, where it is possible to find depressive symptoms below the diagnostic threshold for MDD (but which cause significant distress and impairment in functioning), a condition termed Subthreshold Depression (SD) that is considered on a continuum of severity with MDD. Indeed, SD is often considered to be a precursor of MDD (Wesselhoeft, Pedersen, Mortensen et al., 2015). In childhood and adolescence, it is also common to see a clinical presentation of comorbid anxiety and affective disorders, with some evidence that the former anxiety precedes, and could cause the latter affective disorder (Seligman & Ollendick, 1998). Considering that SD seems to precede MDD and that pre-pubertal onset of MDD leads to worse outcome than onset in adulthood (Van Noorden, Van Fenema, Van der Wee, Zitman & Giltay, 2012), it appears appropriate to develop standardised interventions targeting SD and MDD in childhood and adolescence.

Transactional analysis (TA) is a widely-practiced form of psychotherapy, 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 Trials (RCT) replicated by two independent research groups, or alternatively in at least three Single Case Experimental Design studies (SCED), replicated by at least two (Chambless & Hollon, 1998) or three (Chambless et al., 1998) independent research groups, with each group conducting a case series of a minimum of three cases, without conflicting evidence. 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 such as reverse design (McLeod, 2010). Systematic case study research has been proposed as a viable alternative to RCT and SCED (Iwakabe & Gazzola, 2009). 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 SCED in psychotherapy (McLeod, 2010). HSCED is gaining momentum and enhanced versions have been proposed by different research groups, both to validate new psychotherapeutic approaches and to extend a previously validated psychotherapy to new 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 for conducting sound research according to the available resources within practitioner research networks.

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 case studies as the first step toward recognition of marginalised and emerging models of psychotherapy. 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 et al., 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). These case series have shown that TA can be an effective therapy for MDD 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. Currently, no systematic research has been conducted on TA treatment for depression in childhood and adolescence.

The aim of this study was to investigate the effectiveness of the manualised TA treatment for depression in adults (Widdowson, 2016) for use with adolescents who present with depression. Reviewing the literature, we would expect MDD in adolescence to be characterised by a progression along a continuum beginning with SD and comorbid Generalized Anxiety Disorder (GAD). The present study analyses the treatment of 'Deborah', a 15-year-old Italian young girl who had been suffering from depressive and anxiety symptoms, with a personal history of self-harm, cannabis use and self-induced vomiting, which had emerged in the previous year and had been steadily getting worse in the last few months. The quantitative primary outcomes investigated were depressive and anxious symptomatology; the secondary outcomes were client-generated personal problems and behavioural problems. Quantitative and qualitative analyses were conducted.

## 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 and her parents received an information pack, including a detailed description of the research protocol, and they gave a signed informed consent and written permission to include segments of disguised transcripts of sessions or interviews within scientific articles or conference presentations. They were informed that the therapy would be provided even if they decided not to participate in the research and that they were able to withdraw from the study at any point, without any negative impact on their 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 that does not lead the reader to draw false conclusions related to the described clinical phenomena. Finally, as a member checking procedure, the final article, in Italian language, was presented to the client and her parents, who read the manuscript and confirmed that it was a true and accurate record of the therapy and gave their 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. In this study, the occasional recreational use of marijuana was not considered an exclusion criterion, considering its high diffusion amongst adolescents and that there was no evidence that this habitual use was indicative of a significant physical addiction which would require specialist medical intervention.

### Client

Deborah was a 15-year-old white Italian female adolescent, who lived with her parents and two younger siblings in a large city in north Italy. She was an intelligent, curious, altruistic and resourceful teenager, with many positive values, and who had good self-reflective and evaluative capacities. At the age of 12 she noted some strange symptoms and self-diagnosed (through symptom-checking on the Internet) a metabolic disorder, which was confirmed by subsequent medical investigation. She gained much weight and followed a very strict self-imposed diet, frequently not eating and sometimes inducing vomiting. Deborah attended the second year of high school, with a wide network of relationships but recently had become withdrawn. She had a tendency to easily forgive others, even when they hurt her, and felt "evil and guilty" when she did not. She often self-inflicted injuries on her wrists and ankles, and pulled her hair and picked at her fingernails until they were sore. She had a boyfriend, whom she described as critical and devaluing. She ended this relationship around one month after the onset of therapy. Her mother suffered from depressive disorder. Her father had many episodes of alcohol and drug abuse and was afflicted with a cancer for which he was being treated with a cycle of chemotherapy. Deborah reported having difficulties in her relationship with her parents, especially that she felt unable to communicate with her father. She felt that she was not accepted, not understood, and felt attacked and criticised, and also felt that she was lacking guidance and protection from her parents, and felt strongly guilty for disappointing them.

She independently decided to seek therapy, showing a strong motivation for treatment. In the first appointment, which her parents also attended, the therapist noted that Deborah's parents appeared relieved that they could delegate their daughter's cure to somebody else.

### Therapist

The psychotherapist was a 54-year-old, white, Italian woman with 20 years of clinical experience, and who is a Provisional Teaching and Supervising Transactional Analyst (PTSTA-P). For this case, she received monthly supervision by a Certified Transactional Analyst Trainer (CTA Trainer) with 30 years of experience.

### Intake sessions

The therapy was conducted in a public clinic, once a week and free of charge. The client attended two pre-treatment sessions along with her parents and three individual pre-treatment sessions (0A, 0B, 0C), which were focused on explaining the research project,

obtaining consensus, conducting a diagnostic evaluation according to DSM-5 criteria (American Psychiatric Association, 2013), defining the problems she was seeking help for in therapy along with their duration and severity, developing a case formulation including TA diagnosis, treatment plan and contract, and collecting a stable baseline of self-reported measures for primary (depression and anxiety) and secondary (personal problems and problematic behaviours) outcomes.

Deborah asked to learn how to protect herself, how to express her needs, thoughts and emotions to others, especially her parents, how to regulate her intense and overwhelming emotions and stop self-injuring behaviours.

### DSM-5 Diagnosis

The initial diagnostic phase identified the client's primary diagnosis. Deborah was assessed as meeting DSM 5 diagnostic criteria for moderate MDD: she experienced depressed mood in daily activities for more than one year, most of the day, nearly every day (criterion A1), decreased pleasure in most activities (A2), weight change (A3), restlessness regulated with self-harm (A5) overwhelming feelings of worthlessness and inappropriate guilt (A7), diminished ability to think and concentrate (A8). She also met criteria for moderate GAD: excessive anxiety and worry (criterion A), that were uncontrollable (B), with restlessness (C1), difficulty in concentration (C3) and irritability (C4).

### TA Diagnosis and Case formulation

Case formulation was conducted according the TA diagnostic categories presented in the treatment manual. Deborah assumed a life position (Ernst, 1971; Berne 1972) I'm Not OK, You're OK, that interacted with her stroke economy (Steiner, 1974), which was characterised by an absence of positive strokes and abundance of negative strokes. This in turn led to internalisation of an under-active and under-functioning internal Nurturing Parent and an over-active internal Critical Parent, which activated intense self-critical internal dialogues (Kapur, 1987). Furthermore, the underlying Injunctions (Goulding & Goulding, 1976; McNeel, 2010): Don't be a child (be adult and take care of your parents); Don't think (avoid problem solving) Don't be important (the other's needs are more important than yours), and Don't feel (especially angry when abused) were also identified. These led to the observable drivers (Kahler, 1975) of Try Hard and Please Others and the assumption of drama triangle roles (Karpman, 1968) such as Rescuer with parents and when forgiving friends, Victim when feeling helpless, frail and unable to protect herself, and Persecutor when her active Critical Parent was externalised in conflictual transactions with parents. Script conclusions and decisions (Berne, 1961) were observable through script beliefs and contaminations (Berne, 1961; Stewart & Joines, 1987, 2012) such as: "I am wrong" "others are more important than me", "I cannot be angry with others", "I must forgive and please others", "I must take care of and support my parents",

"my body is not pretty therefore I don't deserve love", "I am evil because I disappoint my parents", "others cannot understand what I mean", "when you have a problem let the time pass away, wait and do not attend to it". The script system (Erskine & Zalcman, 1979; Erskine, 2010) involved all of the above-mentioned thoughts and behavioural manifestations, as well as repressed primary anger when she receives abuse or is not loved and considered by others, which was covered by secondary sadness, helplessness, feelings of being unprotected and unlovable, with worry and restlessness, self-injury behaviours, which in turn triggered the memory recall of episodes of criticism and neglect.

### Treatment

The therapy followed the manualised therapy protocol of Widdowson (2016), including the 12 Key tasks and the research-based principles. Throughout the whole treatment, the therapist focused on 1) building the therapeutic alliance providing empathic listening, 2) giving strong support to the client's self-esteem and recognising her resources and positive strengths; 3) developing the observing self and TA problem solving protocol, in order to enhance Adult functioning, and 4) permeating the sessions with permissions (Crossman, 1966), especially those congruent with the client's injunctions, namely: be and express your needs as adolescent, be important, think (about the consequences on your future), feel and express all your primary emotions, especially anger when you are abused, thus providing systematic implicit experiential disconfirmation and modelling a positive and potent Nurturing Parent. In the first phase (sessions 1-6) the focus was on the recognition and decontamination of script beliefs, emotional literacy and emotion regulation. In the second phase (sessions 7-12) the therapist focused on changing internal dialogue from critical to nurturing and enhancing the client's internal Nurturing Parent. In the third phase (sessions 13-16), the focus was on problem solving strategies, and emotional and behavioural regulation in daily situations.

### 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 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

Four 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  $\geq 10$  has a sensitivity of 88% and a specificity of 88% for major depression (Kroenke, Spitzer, & Williams, 2001) and scores of  $< 10$  are considered subclinical. A change of at least 6 points on PHQ-9 score is considered to assess a reliable improvement or deterioration (RCI).

*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). A change of at least 4 points on GAD-7 score is required in order to assess a reliable improvement or deterioration (RCI).

*Youth-Self Report (YSR)* (Achenbach, 1991; Achenbach and Rescorla, 2001) was rated at session 1, 8, 16, and at 6-month follow-up. YSR is a self-descriptive measure which investigates social competencies and behavioural problems in 11-18 year olds. There are 112 items, coded from 0 (not true in the last six months) to 2 (very often or often true in the last six months). Items measure eight subscale symptoms: Withdrawal, Somatic complaints, Anxiety and depression, Social problems, Thought problems, Attention problems, Aggressive behaviours, and Rule-breaking behaviours (Achenbach, 1991). The first three subscales are grouped into the Internalizing scale, the last two into the Externalizing scale, and the remaining three scales are categorised as Other problems. Overall behavioural and emotional functioning is measured by the Total problems scale. The sum of the scores for each scale and sub-scale may be converted to T-scores for which the manual gives the cut-offs for the clinical and borderline range for boys and girls. A change of at least 6.2 points on YSR Total problems score indicates a reliable improvement or deterioration (RCI).

The *Personal Questionnaire (PQ)* (Elliott, Shapiro, & Mack, 1999; Elliott, Wagner, Sales, Rodgers, Alves & Café, 2016) is 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 3.25 are considered subclinical. In this case series, missing the Italian normative score, for the PQ we adopted a more conservative RCI of two points, rather than the RCI of 1.67 recently proposed by Elliott et al. (2016). The PQ procedure suggests including problems from five areas: symptoms, specific performance or activity (e.g., work), relationships, mood/emotions and self-esteem/internal experience.

All quantitative outcome measures were evaluated according to Reliable and Clinically Significant Change (RCSC) (Jacobson & Truax, 1991). Clinical significance (CS) is obtained when the observed score on an outcome measure drops under a cut-off score that discriminates clinical and non-clinical populations. For example, the PHQ-9 considers a score of  $\geq 10$  as an indicator of current moderate major depression (Kroenke, Spitzer & Williams, 2001). It is important to consider that even under the cut-off score there may be a subclinical disorder. For example, the PHQ-9 considers a score between 0 and 4 an indication of healthy condition, and a score between 5 and 9 as an indicator of mild (subclinical) depression. Reliable Change Index (RCI) is a statistic that enables the determination of the magnitude of change score necessary to consider a

statistically reliable change on an outcome measure (Jacobson and Truax, 1991). In particular, it is helpful in minimising Type I errors which occur when cases with no meaningful symptom change are assumed to have improved. For example, Richards and Borglin (2011) proposed that a reduction of at least 6 points in the PHQ-9 score would be indicative of a reliable improvement. Only when we observe the presence of both CS and RCI, we have a RCSC, which is considered a robust method for assessing recovery in psychological interventions (Evans, Margison & Barkham, 1998; Delgado, McMillan, Leach, Lucock, Gilbody & Wood, 2014). To control experiment error which occurs when multiple significance tests are conducted on change measures, we consider that a RCSC is required in at least two out of three outcome measures, thus demonstrating a Global Reliable Change (GRC) (Elliott, 2015).

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. Deborah's quantitative data are presented in Table 1. Since the client expressed some discomfort in completing self-report measures, we chose to reduce the number of questionnaires the client was required to fill in prior to any session. Thus, we have scores of the GAD-7 only in the assessment phase and at the 1-month and 3-month follow-ups. The first PQ score was available in 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

Deborah's initial depressive score (PHQ-9, 12) indicated a moderate level of depression. The anxiety score (GAD-7, 14) indicated a moderate level of anxiety. The severity score of personal problems (PQ, 4.9) indicated that the client perceived her problems as bothering her somewhere between moderately and considerably. YSR scores show that the sub-scale anxious/depressed symptoms was above the clinical cut off, whereas the sub-scale thought problems was borderline and the other sub-scales were in the normal range. Internalizing problems scale was in the borderline range and externalizing problems scale and total problems scale were in the normal range. Deborah reported in her YSR symptoms such as tics, nervous movements, confusion and difficulties in avoiding negative thoughts, especially severe self-criticism ("I'm wrong").

At session 8, (mid-therapy), depression passed into the subclinical mild range (6), with RCSC. Severity of personal problems decreased to little bothering (3.9). In the YSR, the Anxious/Depressed Thought Problems sub-scales decreased to the normal range whereas the Somatic and Social problems sub-scales increased to the borderline range. All the other sub-scales except that of attention problems increased from values around zero to close to the borderline cut-off. The Internalizing, Externalizing and Total scales were all within the borderline range. Deborah reported in her YSR an increase in alcohol and tobacco consumption and reductions in the use of inappropriate behaviours with adults and peers. At the same time, she described herself as very shy, suspicious, and unable to build intimate relationships with others and she reported episodes of self-harming behaviours. Between the somatic symptoms she reported dizziness, being tired without reason and vomiting.

By the end of the therapy, depression scores passed into the healthy range (3) maintaining RCSC, and personal problems decreased to very little bothering (2.6). Deborah's YSR score revealed a decrease in all the sub-scales except for the Anxious/Depressed one which was in the borderline range and for the Rule-Breaking Behaviour that increased to the clinical range. Similarly, while the Internalizing and Total scales slightly decreased to the bottom of the borderline range, the Externalizing scale increased to the clinical range. Deborah reported that she preferred to spend time with

peers older than herself, that she had stolen things from home and from outside home, that she had used marijuana, and craved for new experiences.

At the 1-month follow-up, depression scores remained in the healthy range (3) with RCSC. Anxiety passed to subclinical mild range (8), with both clinical significance and reliable improvement (RCSC) compared to assessment. Personal problems returned to moderately bothering (4.3) without reliable deterioration.

At the 3-month follow-up, depression remained in the healthy range (3), anxiety remained in the mild range (5), and personal problems passed to be considered as not bothering at all (1.3), with RCSC.

Finally, at the 6-month follow-up depression returned to the mild range maintaining RCSC (5), personal problems returned to very little bothering, maintaining RCSC (2.3). YSR scores indicated a global increase in all the symptoms except for the Social problems sub-scale. In particular, the Rule-Breaking Behaviour and Anxious/Depressed sub-scales and all global scales were in the clinical range. A great number of items obtained the maximum score of 2 with the co-presence of anxious behaviours such as being nervous and worried, excessive crying and eating, and deviant behaviours such as the use of drugs, disregarding social and parents' rules and bullying behaviours.

Table 2 shows the 10 problems that the client identified in her PQ at the beginning of the therapy and their duration. Problems are related to: symptoms (1, think; 3, hurt), specific performance or activity (4, tell; 6 stop, 9 be understood), relationships (7 attacked), mood (5, disappointing, 8, uncontainable) and self-esteem (2, fattening; 10 unaccepted). Two problems were rated as bothering her maximum possible, two as very considerably, two were rated considerably bothering, two as moderately. Two problems were rated under the clinical cut off, and hence without clinical significance. Three problems lasted from 3-5 years, two from 1-2 years, four from 6-11 months and one from 1-5 months, representing an almost stable and longstanding baseline. The longer lasting problems were related to relationships, mood and self-esteem, where symptoms were present for less than two years.

At the end of the therapy 5 out of the 8 problems above the clinical cut off showed a RCSC, and 2 showed a reliable improvement (RCI). At the first follow-up almost all items reliably deteriorated. At the 3-month follow-up, all the problems above the clinical cut off showed a RCSC, and four maintained RCSC at the 6-month follow-up. Her symptoms (items 1 and 3) showed the larger and more stable change.

Figures 1 to 4 below allow time series' visual inspection of the weekly scores of primary (PHQ9 and GAD-7) and secondary (PQ and YSR) outcome measures. Concerning the YSR we report in Figure 3 the sub-scales that evidenced behavioural symptoms within the borderline or clinical range during the therapy

	Pre-Therapy	Session 8 Middle	Session 16 End	1 month FU	3 months FU	6 months FU
PHQ-9	<b>12</b> Moderate	6 (+)(*) Mild	3 (+)(*) Healthy	3 (+)(*) Healthy	3 (+)(*) Healthy	5 (+)(*) Mild
GAD-7	<b>14</b> Moderate	-	-	8 (+)(*) Mild	5 (+)(*) Mild	-
PQ <sup>a</sup>	<b>4.9<sup>a</sup></b> Moderately	<b>3.9</b> Little	2.6 (+) Very little	<b>4.3</b> Moderately	1.3 (+)(*) Not at all	2.3 (+)(*) Very little
YSR: Total Problems <sup>b</sup>	57 Normal	63 Borderline	61 Borderline			<b>67</b> Clinical
Internalizing Problems <sup>b</sup>	61 Borderline	64 Borderline	60 Borderline			<b>67</b> Clinical
Externalizing Problems <sup>b</sup>	49 Normal	63 Borderline	<b>66</b> Clinical			<b>70</b> Clinical
-Anxious/ Depressed <sup>b</sup>	<b>70</b> Clinical	61 Normal	65 Borderline			<b>72</b> Clinical
-Withdrawn/ Depressed <sup>b</sup>	52 Normal	60 Normal	50 Normal			59 Normal
-Somatic Complaints <sup>b</sup>	52 Normal	65 Borderline	59 Normal			62 Normal
-Social Problems <sup>b</sup>	52 Normal	64 Borderline	51 Normal			51 Normal
-Thought Problems <sup>b</sup>	67 Borderline	52 Normal	56 Normal			61 Normal
-Attention Problems <sup>b</sup>	54 Normal	52 Normal	54 Normal			64 Borderline
-Rule-Breaking Behaviour <sup>b</sup>	51 Normal	63 Normal	<b>72</b> Clinical			<b>74</b> Clinical
-Aggressive Behaviour <sup>b</sup>	52 Normal	62 Normal	54 Normal			62 Normal

*Table 1: Deborah's Quantitative Outcome Measure*

**Note.** Values in **bold** are within the clinical range; + indicates clinically significant change (CS). \* indicates reliable change (RC). 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). YSR = Youth-Self Report (Achenbach, 1991). FU = follow-up.

Clinical cut-off points: PHQ-9  $\geq 10$ ; GAD-7  $\geq 10$ ; PQ  $\geq 3.25$ ; YSR symptoms scales  $>69$ ; YSR cumulative scales  $>64$ . Reliable Change Index values: PHQ-9 improvement of six points, GAD-7 improvement of four points, PQ improvement of two points, YSR change of 6.

<sup>a</sup>First available score in session 1. <sup>b</sup>YSR is expressed in T-scores



	PQ items	Duration	Pre-Therapy <sup>a</sup>	Session 8 (middle)	Session 16 (end)	1 month FU	3 months FU	6 months FU
1	I cannot succeed in lingering and think about things	1-2y	<b>7</b> Maximum possible	<b>5 (*)</b> Considerably	2 (+)(*) Very little	1 (+)(*) Not at all	1 (+)(*) Not at all	1 (+)(*) Not at all
2	"Fattening up" makes me think I have less attentions	3-5y	<b>7</b> Maximum possible	<b>5 (*)</b> Considerably	<b>4 (*)</b> Moderately	<b>7</b> Maximum possible	1 (+)(*) Not at all	<b>5 (*)</b> Considerably
3	When I lose control, I hurt myself	6-11m	<b>6</b> Very considerably	<b>4 (*)</b> Moderately	1 (+)(*) Not at all	<b>6</b> Very considerably	1 (+)(*) Not at all	1 (+)(*) Not at all
4	I cannot succeed in telling my dad what I think	1-2y	<b>6</b> Very considerably	<b>4 (*)</b> Moderately	<b>4 (*)</b> Moderately	<b>6</b> Very considerably	2 (+)(*) Very little	3 (+)(*) Little
5	I feel I have disappointed my parents	6-11m	<b>5</b> Considerably	<b>7</b> Maximum possible	3 (+)(*) Little	<b>5</b> Considerably	2 (+)(*) Very little	<b>4</b> Moderately
6	I feel I'm not protecting myself and I let others use me	6-11m	<b>5</b> Considerably	2 (+)(*) Very little	2 (+)(*) Very little	<b>5</b> Considerably	1 (+)(*) Not at all	<b>4</b> Moderately
7	I cannot succeed in accepting the reactions my dad has with me and I feel attacked	3-5y	<b>4</b> Moderately	<b>4</b> Moderately	<b>4</b> Moderately	<b>4</b> Moderately	2 (+)(*) Very little	3 (+) Little
8	I don't succeed in containing my emotions	3-5y	<b>4</b> Moderately	3 (+) Little	2 (+)(*) Very little	<b>4</b> Moderately	1 (+)(*) Not at all	2 (+)(*) Very little
9	I don't succeed in making myself understandable (I cannot accept that others won't understand me)	6-11m	3 Little	3 Little	2 Very little	3 Little	1 (*) Not at all	1 (*) Not at all
10	I do not feel accepted by my relatives for my metabolic disease	1-5m	2 Very little	2 Very little	2 Very little	2 Very little	1 Not at all	1 Not at all
	Total		49	39	26	43	13	25
	Mean		<b>4.9</b>	<b>3.9</b>	2.6	<b>4.3</b>	1.3	2.5

Table 2: Deborah'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  $\geq 3$ . 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.

<sup>a</sup>The first available score was in session 1.

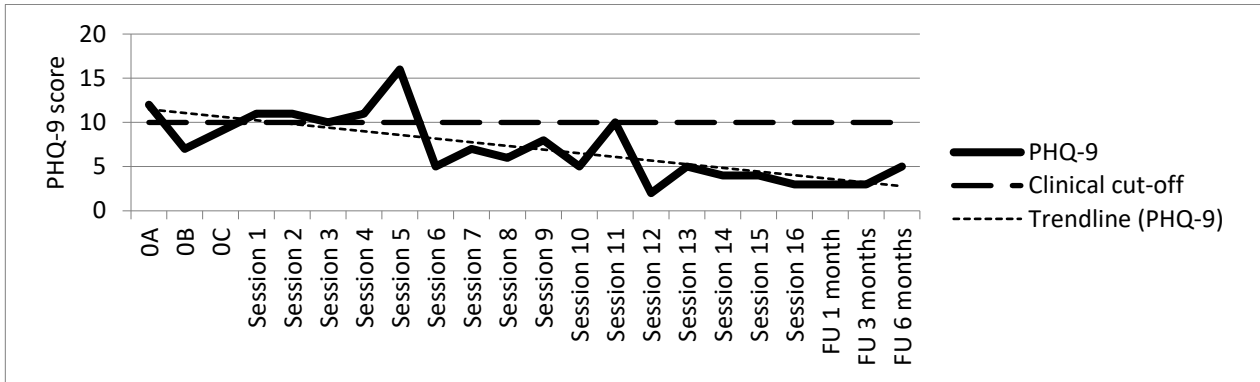


Figure 1: Deborah's weekly depressive (PHQ-9) score

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

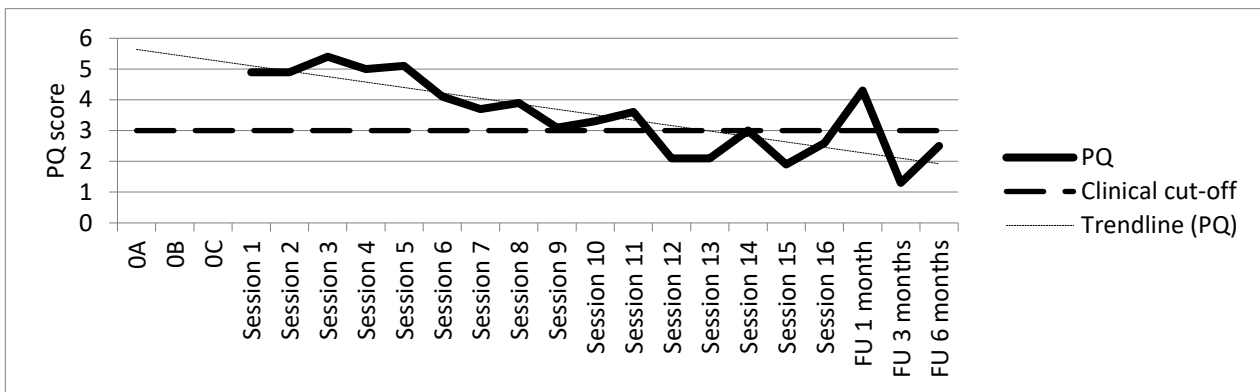


Figure 2: Deborah's weekly personal problems (PQ) score

**Note.** The first available score was in session 1. 0A, 0B and 0C = assessment sessions. FU = follow-up. PQ = Personal Questionnaire (Elliott, Shapiro, & Mack, 1999).

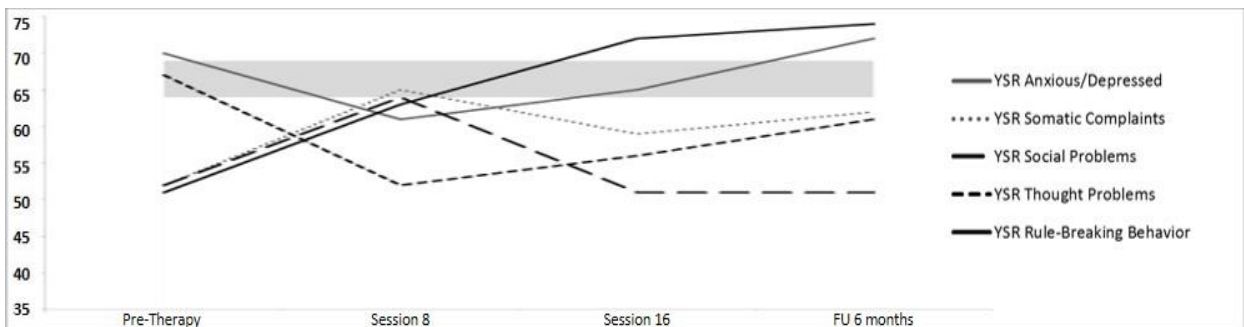


Figure 3: Deborah's Youth Self Report (YSR) sub-scales T-scores

**Note.** FU = follow-up. YSR = Youth-Self Report (Achenbach, 1991). The grey area indicates the borderline range. Above the grey area scores are considered in the clinical range, below in the normal range.

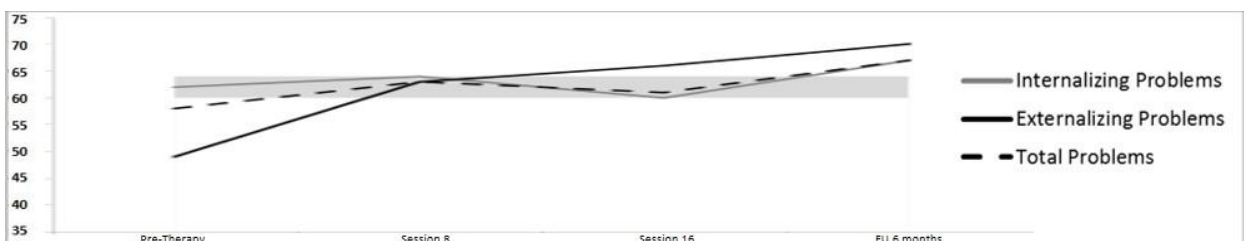


Figure 4: Deborah's Youth Self Report (YSR) scales T-scores

**Note.** FU = follow-up. YSR = Youth-Self Report (Achenbach, 1991). The grey area indicates the borderline range. Above the grey area scores are considered in the clinical range, below in the normal range.

Session	Rating	Events	What made this event helpful/important
1	8 (greatly)	Talking about things that I normally keep to myself	Understood that I don't have to feel guilty for problems that are not mine and that don't concern me
2	8 (greatly)	Not crying while talking about my cuts. Having connected the cuts not to the pain but to anger/cruelty with myself	Not getting angry with myself if I'm not involved
3	8.5 (greatly)	Understanding that many people don't deserve me	It has been useful because I realized something I have never considered
4	7.5 (moderately)	When the psychologist told me "two kilos won't make any difference"	Understanding that two kilos are not important
5	7 (moderately)	Not feeling less than my grandmother	It's important because I understood that we are on the same level
6	8 (greatly)	"I'm not the Parent, I'm the Child"	Think more about myself
7	7 (moderately)	There are two parts of me: a positive and a negative one	I accept only the positive one
8	8 (greatly)	Worrying more about myself	Realizing that I have to worry a little bit more about myself
9	8 (greatly)	I try to protect myself a little bit more	Put boundaries or rules
10	9 (extremely)	"It seems that your mother wants to be supported by you"	I'm not the parent
11	9 (extremely)	"Your parents are behaving like parents"	My relationship with them has changed
12	9 (extremely)	"You are intelligent and 'abundant', there is no need to do drugs"	I felt right and "special"
13	9 (extremely)	I have to find some boundaries for myself	-
14	9 (extremely)	"I feel more comfortable with myself"	I have more self-esteem
15	9 (extremely)	"I don't feel ready to make my parents feel guilty"	I still think about protecting myself
16	9 (extremely)	"All problems are set aside when you smoke!"	Maybe I have an addiction

**Table 3: Deborah'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).

### Qualitative Data

Deborah compiled the HAT form at the end of every session (Table 3), reporting positive/helpful events and one hindering event. The hindering event was reported in session 6 and rated 3 (moderately hindering): "My family lack of Parents", suggesting a misunderstanding of the meaning of in-therapy hindering events. All positive events were rated from 7 (moderately helpful) to 9 (extremely helpful). She reported helpful aspects on: symptoms (1 guilty, 2 hurt, 16 addiction); relationships (6 I'm the child, 9 boundaries, 10 support, 11 change); mood/emotion (2 angry); self-esteem/inner experience (3 deserve, 5 feeling less, 7 accept, 8 worry about me, 12

special, 13 boundaries, 14 comfortable, 15 protect myself).

Deborah participated in a Change Interview 1-month after the conclusion of the therapy. In this interview she identified her main and significant changes (see Table 4). Deborah described her therapy as "a place where I could talk to a friend... it was a liberation, like emptying everything... It has been nice" (Client line 12), "because I succeeded in talking about things I normally don't speak of" (C6). "When I got back home I felt lighter" (C112), "it made me reach some goals that without a psychologist I would have never reached" (C5). When Deborah started

the therapy, she used to hurt herself because she “needed to feel external pain, in order to stop feeling the pain inside” (C33), whereas now she is “not doing this anymore. I try to focus on other things, like school, I go out with my friends” (C33). At the beginning of the therapy, she used to think she was disappointing her parents, whereas now she does not think about it anymore (C33). Deborah in her CI did report two negative, obstructive or unpleasant aspects of therapy: firstly, that she found filling in the outcome measures as “boring” (C121) and secondly that she was “thinking too much” as a consequence of therapy (C52).

Four changes reported by Deborah are related to her ability to think and solve problems (items 1, 3, 6 and 11). Three changes are related to her initial symptoms: hurt herself (2), feeling guilt (4) and drug use (5). Other change refers to self-protection (7), self-esteem (9) and differentiation from parents (10). All changes are rated

very or extremely important for her. Three changes are rated unlikely, five somewhat unlikely, one neither and two somewhat likely without therapy. Two changes were surprising, four somewhat surprising, two neither, one somewhat expected and two expected. According to Deborah, all these improvements happened because she “worked hard and found a very attentive therapist, who listened, understood and remembered things I told her, and also because back home you can think about what you said, what happened during the session” (C100).

Deborah also reported that her mother used to check her body each week, to ensure that she was not cutting herself, “but now I understand, and she understands that I’m different, that something has changed, and that I have changed, she is not checking my body any longer” (C59). The client also stated that therapy had been useful in helping her to change her way of thinking (C107).

Change	How much expected change was <sup>(a)</sup>	How likely change would have been without therapy <sup>(b)</sup>	Importance of change <sup>(c)</sup>
Standing in front of problems, I succeed in finding a way out	3 (neither)	2 (somewhat unlikely)	4 (very)
When I suffer, I don’t hurt myself anymore	3 (neither)	2 (somewhat unlikely)	4 (very)
I’m able to think about things that make me feel bad	4 (somewhat surprised)	4 (somewhat likely)	5 (extremely)
I don’t think about disappointing my parents anymore	5 (surprised)	2 (somewhat unlikely)	5 (extremely)
I stopped taking drugs	4 (somewhat surprised)	2 (somewhat unlikely)	4 (very)
I succeed in lingering, pausing and thinking about consequences	2 (somewhat expected)	1 (unlikely)	5 (extremely)
I protect myself, I think about myself	5 (surprised)	2 (somewhat unlikely)	4 (very)
I cry less	1 (expected)	1 (unlikely)	5 (extremely)
I accept myself for what I am	4 (somewhat surprised)	1 (unlikely)	5 (extremely)
I accept my father for being different from me	4 (somewhat surprised)	4 (somewhat likely)	5 (extremely)
I think more	1 (expected)	3 (neither)	4 (very)

Table 4: Deborah’s Changes identified In the Change Interview

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

<sup>a</sup>The rating is on a scale from 1 to 5; 1= expected, 3 = neither, 5 = surprising.

<sup>b</sup>The rating is on a scale from 1 to 5; 1=unlikely, 3 = neither, 5 = likely.

<sup>c</sup>The rating is on a scale from 1 to 5; 1 = slightly, 3 = moderately, 5 = extremely.



## HSCED Analysis

### *Affirmative Case*

The affirmative team identified four lines of evidence supporting the claim that Deborah 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 depressive symptomatology measured with PHQ-9; visual inspection shows a clinically significant improvement since session 6, reliable since session 12, that is maintained until the end of the treatment and throughout the follow-up intervals. Anxious symptomatology, measured with GAD-7 in the phase 0 and then again at the first and second follow-up, shows a reliable and clinically significant improvement. YSR total problem scale was in the non-clinical range at the beginning of the treatment and showed a global, not reliable and not clinical deterioration during the treatment, that became reliable and clinically significant at the 6-month follow-up. The anxiety/depressive subscale was in the clinical range at the beginning of the treatment, and showed an early, clinically significant and reliable improvement in session 8, and in following assessment showed a deterioration, however not reliable in respect the pre-treatment score. In the PQ (Table 2), Deborah identified 10 main problems at the beginning of the therapy that she was trying to solve. All problems reached reliable and clinically significant improvement by the second follow-up (except item 9 and 10 that were in the subclinical range at pre-treatment assessment) and four maintained the reliable and clinically significant improvement at the six-month follow-up. All the stable problems, lasting from 1 to 5 years, showed a clinical and/or reliable change at the 6-month follow-up. Overall, there is support for claiming of a global reliable change.

Qualitative data supports these changes in stable problems. In her Change Interview (CI) Deborah says "I reached my goal, so I did it, that's it" (CI, C5) and reports as a main achievement to be able to think about the consequences of her actions (CI, C46-47): "I decided to seek therapy for what I was doing to myself, for how I treated myself. I never thought of what could happen next. Now it's different, I tell myself 'what are you doing? Wake up!', whereas before I was not able to think like this. My head changed, I changed how I think". As for her habit of smoking marijuana, Deborah said "If friends offer it to me, I still have a 'toke' [inhale], but it's not like before, absolutely, and therapy helped me in this. I have to thank her (the therapist), because if I hadn't continued the therapy, I wouldn't have understood many things, and I would probably still be addicted, I'm sure of that" (C39).

#### *Retrospective attribution*

Deborah identified in her Change Interview eleven important changes, eight of them rated unlikely or somewhat unlikely without therapy (Table 4). She recognised that the therapy allowed her to change her way of thinking (CI, C107-111), to stop feeling she was disappointing her parents (C48), to develop a major capacity to control her emotions instead of crying (C60),

to stop using marijuana (C39), and especially to stop hurting herself (C40). Deborah referred that she liked talking to her therapist: "in less than five sessions, I felt like she knew everything about me, as if she was my sister, my best friend, because she knew everything. It felt good, because I succeeded in talking about things, which I normally do not do with others. It helped me to feel free to tell her my problems, to trust her" (C6-8). "Some aspects of my life changed drastically, before, when I woke up in the morning I thought 'Here we are, another day to deal with', whereas now I wake up and I feel normal" (C10). A very helpful aspect of therapy for Deborah was "feeling free from judgements" (C101). She also affirmed that there were no negative aspects, obstacles or unhelpful aspects to her therapy.

#### *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 Deborah found helpful in each session. All reported events (but one) are considered moderately to extremely useful and are coherent with the diagnosis, the treatment plan and the interventions reported in the therapist's notes. Changes in depression and anxiety symptoms (Table 1), and in particular, feelings of guilt and worry, and personal problems (Table 2) appear tied to interventions on changing her internal dialogue from Critical Parent to Nurturing Parent (HAT, Table 4, sessions 1, 6, 8), self-esteem (HAT, Table 4, sessions 4, 5, 11, 12, 14) and problem solving (HAT, Table 4, sessions 9, 13, 15). Changes in her drug use (CI, Table 4, item 5) appear associated to improved comprehension of the role of the drug and discovering alternative ways to manage her needs and emotions (HAT, Table 3, session 16). Changes in self-injury behaviour (CI, Table 4, item 2) appear tied to the growing awareness and comprehension that she tended to cut herself, starve herself, self-induce vomiting, and made use of alcohol and drugs in order to punish herself for thinking that she had disappointed her parents (HAT, Table 3, session 2). Changes related to thinking and problem solving (CI, Table 4, item 1, 3, 6, 11) appear tied to intervention of the problem-solving protocol (HAT, Table 3, session 1, 3, 8, 9, 15).

#### *Event-shift sequences (process to outcome mapping)*

The greater effect on depressive symptoms appeared to be tied to interventions in first sessions, on changing the self-critical internal dialogue associated with her feelings of guilt and worthlessness. Interventions on changing this internal dialogue and self-esteem are mirrored in HAT forms (session 1, 2, 4, 5) and reflected in subsequent changes in PHQ-9 scores on item 2 feeling down, depressed and item 6 feeling bad about yourself. Since the beginning of therapy, the therapist focused on Deborah's self-injury, connecting it with unexpressed anger (Table 3, HAT session 2), that led to an improvement in her PQ scores (Table 2, item 3), which was maintained at the six-month follow-up. In session 3, the therapist worked on decontaminating script beliefs about Deborah's way of relating to others. Deborah

realised she had adopted her mother's model of relating, responding passively and powerlessly to her boyfriend, in the same way her mother did with her own husband. Deborah recognised that she was used to avoiding problems and not thinking about solution, instead of actively problem-solving. This was reflected in following changes in PQ items 1, 8, 9. From session 6, the therapist worked on restoring the roles of daughter and parents. This is reflected on HAT 6 ("I'm not the Parent, I'm the Child") and led to a change in role-assumption and her relationship with her parents that is reflected in HAT 11 ("parents are behaving like parents", "My relationship with them has changed").

### **Sceptic Case**

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

Three of the quantitative measures used (PHQ-9, GAD-7 and PQ) are not validated for adolescence, and should therefore not be used in this case study. The only measure validated for adolescents is the YSR, which at the 6-month follow-up showed a reliable deterioration in all cumulative scales (internalizing, externalizing, total problems). Anxiety/depression subscale indicated a non-reliable improvement at the end of the therapy, and a deterioration worse than the pre-therapy score at the 6-month follow-up. For such reasons, we reject the claim of a global reliable change. Indeed, the deterioration observed in all cumulative scales at session 8, at the end of the therapy and at the follow-up, and the deterioration observed on the rule-breaking sub-scale, suggest a negative impact of the therapy. Also, in qualitative data we note evidence of inconsistent change: in the CI Deborah tells about having stopped using drugs, yet in the 6-month follow-up she indicated in the YSR, at item 105 I use drugs for non-medical purposes (this item did not include alcohol or tobacco) she scored 2, very true/often true.

*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.*

Even considering the PHQ-9 as a valid measure for depression in adolescence, the visual inspection of the three-point baseline shows an unstable pattern, making it difficult to calculate a reliable change. The PQ scores have not been collected in the pre-treatment assessment, so the case lacked a stable baseline to examine subsequent scores.

*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, Deborah reported almost only positive comments about the therapy and the therapist, and in her HAT forms she reported only one hindering event (that is also unclear since it appeared to refer to extra-therapy events). She described the therapist as if she were "her

best friend" (C6). For this reason, it is possible that Deborah's tendency to 'compliance' might have affected her outcome measures. Furthermore, in almost all sessions, she showed compliance to the therapist's interventions.

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

Deborah searched independently and spontaneously for therapy, and three out of eleven changes reported in her CI were expected or somewhat expected, and two neither expected nor unexpected. This suggests that the change can be partially tied to self-persuasion and personal expectancy of a resolution of her problems.

*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.*

Even accepting data from PHQ-9 and GAD-7, changes and widely fluctuating scores are normal in adolescence and all observed changes can be attributed to normal fluctuations associated with adolescence. Furthermore, Deborah discovered that she was suffering from a metabolic disorder the previous year. In fact, two problems out of ten on her PQ concern her illness (Table 2, items 2 and 10). Changes in depression and anxiety symptoms may represent a normal reaction to her illness, followed by a return to the previous condition.

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

The improvements in Deborah's quantitative data may have been due to extra-therapeutic events, such as: her father's long absence for work, might have led to a reduction in conflicts (session 6, C4; and 7, C5); the end of the relationship with her devaluing boyfriend (session 5, C 23); meeting a new boy she liked (session 7, C5); and finally, participation in a support group for adolescents with the same illness (session 11, C3).

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

Depression may be tied to Deborah's metabolic illness and change in depression may be tied to hormonal balance following pharmacological treatment. The use of marijuana might also have had some effect on her mood, making it difficult to differentiate between effect of the therapy and effect of the biological processes.

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

Participating in the research, talking about her problems and being recorded made Deborah feel embarrassed (CI, C4), thinking "what if I swear, I must be very careful" (CI, C5), which might have affected quantitative and qualitative data. She also found outcome measures very boring (CI, C122), suggesting a possible inaccurate compilation of the forms.

### **Affirmative Rebuttal**

1) A search for existing literature indicated that there are several studies which support the use of PHQ-9 and GAD-7 with adolescents. Studies indicate that disorder sensitivity and specificity in adolescence are similar to those of adult population, suggesting only a slightly higher clinical cut off. Even if there is not a validated version for Italian adolescents, there is no reason to suppose a different result might occur in light of the other validations. Thus, we affirm the presence of a global reliable change. The global deterioration observed in YSR may reflect the progressive trust of the client towards the therapist and consequent self-disclosure in reporting her problems accurately. The items of the YSR are defined as concrete behaviours, and in the pre-therapy assessment the client may have had some hesitation in describing herself, in fact she signed the item "I use drugs for non-medical purpose (don't include alcohol or tobacco)" with 0 (never) however in following sessions she stated that she used marijuana with friends.

2) Fluctuation in the PHQ-9 scores in the pre-treatment phase are inferior to the reliable change index, thus are not reliable and may reflect the error measure of the test. The PQ is a client-generated measure that is built with the client during the pre-treatment phase, so the first score is usually available only at the first session. The stable baseline is supported by the duration form that assesses how long the personal problems were bothering the client.

3) The creation of a friendly relationship is considered a necessity to engage adolescents in therapy. In her CI, the client appears able to describe unpleasant aspects of the therapy, such as boredom in filling the forms (CI, C122) and freely spoke about some goals that had not been reached. For example, regarding the relationship with her father, she admitted she hoped to change it, but during the course of therapy she realised that it was not possible: "this did not change, of course it didn't, I somewhat expected it... but this is who he is, he did not take part in these sessions, so my head changed, not his, and I accept it" (C55), suggesting that the relationship was friendly but not compliant.

4) Most of the change reported by the client in her CI are rated as unexpected or somewhat unexpected, suggesting that it is not present as a personal expectancy artefact.

5) When Deborah presented for therapy, her condition was worsening, to the point she felt the need for help. Despite the typical fluctuating patterns of adolescents, the client identified in her PQ and CI different changes in problems that were not resolved in previous years by the simple passage of time or natural course of the disorder, which contradicts a 'reverse to normal baseline' hypothesis.

6) The observed extra-therapeutic events are normal events in the life of an adolescent and in the change interview the client did not refer to them as a cause of her

change. The meeting with the new boy is not followed with a stable relationship, and happens after the wider improvement in depressive symptomatology observed in session 6. Furthermore, the client described a wide network of friendships at the beginning of the therapy, and it appears improbable that the new friendship group would substantially affect her depression.

7) The use of marijuana and pharmacotherapy began largely before the worsening that led the client to the therapy, thus suggesting the absence of a causal role in observed improvement.

8) The overall transcriptions of the sessions show that Deborah expressed without censoring herself, furthermore in her CI Deborah reported feeling initially embarrassed, "then I got the hang of it, I got used to it, I don't even realise it" (C5), thus supporting the claim that the client was not influenced by the participation in research. As for boredom, quantitative measures, Deborah reported also "they were very boring, but then you see how you have changed... when you conclude therapy you say 'Crap! I reached this goal, I made it!' and I think that without these evaluations you cannot realise it" (CI, C122), suggesting an adequate compliance in filling in tests and forms.

### **Sceptic Rebuttal**

Even accepting the use of outcome measures not validated for adolescence, there still remain the difficulties in the use of questionnaires for adolescents. There are several indications that Deborah found some difficulties, inconsistency and confusion in completing the questionnaires. For example, during the CI, when reviewing the PQ, Deborah noticed that she made a mistake in rating the item 10 in the first session: she stated "'very little'... it was supposed to be the opposite, I don't know what happened, it was supposed to be 'maximum possible', before I didn't think, whereas now it's 'not at all'" (CI, C148). She also misunderstood how to rate how likely the change would have been without therapy, scoring the attribution by inverting the scale: somewhat unlikely without therapy... "I believe therapy helped me a lot, so 4 (somewhat likely)" (C68), and repeating this procedure to be sure she understood how to rate it correctly she asked "without therapy, what did I say?", the interviewer replied "4", "yes, somewhat unlikely, that's correct" (C76-77). Furthermore, in the CI Deborah referred "I have always thought that I was a disappointment for my parents" (C33), whereas, in her PQ's duration form she wrote it lasted only from 6-11 months. In her HAT, she described as a hindering in-therapy event that in her family there was a lack of clarity regarding who would take parental roles. Overall, this suggests that scores on quantitative data could reflect her difficulties in understanding the measure and in quantifying improvements in different problems.

### **Affirmative Conclusion**

Deborah's depression, anxiety, personal and behavioural problems were related to difficulties in sustaining a self-nurturing internal dialogue, harsh self-



criticism and difficulty in solving problems. The therapist created a warm relationship where the client felt free to be open and experienced strong support for her compromised self-esteem. The focus on awareness of her internal dialogue, differentiation between internal dialogues from Critical and Nurturing Parent, the ability to think about herself and her future, the systematic engagement in problem solving activity and the systematic provision of permission contrasting her injunctions, led to a stable change in depressive and anxiety symptoms, especially guilt and worthlessness. This also resulted in a behavioural change in self-injury and marijuana use, and a change in interpersonal relationships, where she now expresses her anger and puts boundaries in place with others. This change was reflected also in the adoption of new roles and rules within her family, with the client accepting her role of daughter rather than parent.

### *Sceptic conclusion*

Deborah's symptoms arose after discovering her metabolic illness that made her gain 10 kilos, just while entering adolescence, a phase of life generally difficult and abundant in different problems. Some quantitative measures are not validated for adolescents and the tests also present several errors in their compilation. Several extra-therapeutic events may have had a prominent role in the reversal of symptomatology. The observed change could be due to a spontaneous remission.

### **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 is a clearly good outcome (60% of certainty) with aspect of a mixed outcome (40% of certainty)

Quantitative data show a reliable and clinically significant change on measures of depression (PHQ9), anxiety (GAD7) and Personal problems (PQ), maintained at the follow-ups, allowing a claim for a global reliable change. Qualitative data supports claims of change in self-injury, marijuana use, self-esteem and problem-solving ability, all of which are changes that the client attributes to the therapy. Aspects of mixed outcome are mainly connected to the YSR scores, which at the 6-month follow-up indicated a global increase in most problem behaviour dimensions.

**Judge B (SM).** This is a clearly good outcome (80% certainty) or a mixed outcome (20%). Quantitative and qualitative data described in affirmative brief and rebuttal appear to provide clear support for a claim that this was

a good outcome case. Aspects of mixed outcome are related to the sceptic briefs and rebuttal, that suggest a more conservative position.

**Judge C (AP).** This case is classifiable as good outcome case (80% certainty) to mixed case (20% certainty). Quantitative data suggests a reliable and global change, partially maintained at the six-month follow-up. Behavioural change in self-injury, use of marijuana, and problem-solving ability appear to be the most important achievements for this client, supported by session transcripts and Change Interview. The client became more differentiated from her father by ending her controlling behaviours towards him and understanding that they are two different persons. At the same time, she began to ask her parents to fulfil their role. These kinds of change are not only tied to a symptomatic remission, but appear tied to a deeper change in thinking and self-esteem.

### *Opinions about the degree of change*

**Judge A.** The client's change is considerably (60%, with 80% of certainty), considering the global reliable change and the importance of change in self-injury behaviours. The change in quantitative measures appears stable during the follow-ups. Moreover, qualitative data (Change Interview) shows a consistent change in Deborah's self-esteem, (disappointing her parents vs no more disappointing, not protecting herself vs protecting herself), indicating a change in deep dimensions of personality and not merely symptomatic changes.

**Judge B.** The client changed substantially (80% with an 80% certainty). There is strong evidence that Deborah has more self-esteem, she stopped hurting herself, she likes herself more, she understands her emotions and feelings, she does not lose control anymore, she protects herself and she is able to think about the consequences of her actions. This shows a significant change. Furthermore, in her CI Deborah reports not feeling that she is a disappointment for her parents anymore, being able to think about negative things and understanding that making use of drugs was not a solution for feeling better and drawing her parents' attention.

**Judge C.** The client showed a substantial change (80% with 80% of certainty). The client's quantitative data (clinical significance, reliable change index and stability during follow-up) showed a clear improvement, especially in her depressive symptomatology. Furthermore, qualitative measurements represent a full description of Deborah's improvement, traceable in the Change Interview and in relational episodes described in her last sessions. These aspects support a widely significant change in her self-representation: she seems to feel more certain of herself, to have more self-esteem, to feel more comfortable with her body, less worried about taking care of her parents and more focused on taking care of herself, and she ceased self-injuring behaviours and taking drugs.



*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. Qualitative data in the HAT form (summarised in Table 3) of the client is extremely helpful to understand what the client felt important in the course of therapy, such as increasing self-esteem and stop hurting herself in many different ways. Qualitative data in the Change Interview (summarised in Table 4) reported a list of changes attributed to the therapy, for example the client explicitly recognised that without the therapy she would still be addicted. There were clear connections between interventions, HAT forms and changes drawn in the affirmative case which appear credible.

**Judge B.** The change is substantially (80% with 80% of certainty) due to the therapy. Since the beginning of therapy, the therapist focused on Deborah's tendency to think she was not doing enough to be a good child and on her difficulty in appreciating and loving herself. All sessions involved deep work on her self-esteem, also reported in Deborah's HAT forms (Table 3: session 1, 4, 5, 7, 8, 9, 12 and 14). Furthermore, in the CI she reported that eight out of eleven big changes were due to therapy (Table 4).

**Judge C.** The change appears substantially due to the therapy (80% with 80% of certainty). One of the most important changes is that the client gave up taking care of her parents, which had previously created a frustrating situation where she was the parent of her own parents. This frustrating situation appeared to be tied to depressive symptoms and negative self-representation.

The therapist focused on helping the client understand she is not responsible for her parents, and can ask them to protect her as a young girl of her age requires.

Furthermore, the therapist helped the client to focus on her needs and to express them. These aspects can be found in her HAT forms (Table 3) and CI, where she described how she began to feel like she is a child rather than a parent (HAT 6). This produced a change in the relationship with her parents, a decrease in her frustration, and as a result a decrease in her self-injuring behaviours and her use of drugs.

*Mediator Factors*

**Judge A.** A good therapeutic alliance and empathic listening were essential in order to reinforce self-esteem. During the therapy self-injuring behaviours were connected to depressive feelings, beneath which there were: anger toward her parents, need for boundaries from parents and holding of emotions, and feelings of having disappointed her parents. Changes in self-criticism and problem solving strategies appear to be at the base of the change process.

**Judge B.** A good therapeutic alliance has been essential for exploring Deborah's worries, emotions and feelings in order to work on her depressive symptoms and to have more self-esteem. Also internal dialogues which generated feelings of guilt and low self-esteem have been successfully analysed.

**Judge C.** The most significant mediator factor seems to be the ability to differentiate between the role of child and parent, to feel her own needs, recognise they are allowed and experience permission to express them. Furthermore, the therapist succeeded early in the therapy in establishing a good therapeutic alliance. The therapist acted as a model of affective, protective and Nurturing Parent, allowing the client to have a new them in establishing a good therapeutic alliance. The positive experience and change her internal dialogue which increased her self-esteem.

	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?	60%	80%	80%	73.3%
To what extent did the client change over the course of therapy?	60% Considerably	80% Substantially	80% Substantially	73,3% Considerably to Substantially
How certain are you?	80%	80%	80%	80%
To what extent is this change due to therapy?	80% Substantially	80% Substantially	80% Substantially	80% Substantially
How certain are you?	80%	80%	80%	80%

*Table 5: Adjudication results.*

### *Moderator Factors*

**Judge A.** Being very clever and smart, and having a wide network of relationship with peers, have been important resources for the client in order to make the best use of this therapy, which she sought spontaneously.

**Judge B.** The client was very motivated since the beginning; that and her cleverness were both surely powerful resources that helped Deborah make the best use of the therapy.

**Judge C.** The client showed a positive attitude toward the therapist and therapy and was very motivated to work on herself during therapy. She was also extremely intelligent and introspective.

## Discussion

This case aimed to investigate the effectiveness of a manualised TA treatment for depression in adulthood in an adolescent with moderate MDD in comorbidity with GAD, self-injury and substance use. Primary outcomes were depressive and anxiety symptomatology, that showed a reliable and clinically significant change, maintained throughout follow-ups. Secondary outcomes were personal problem and behavioural problems, that showed a mixed outcome. The therapist conducted the treatment in a good to excellent adherence to the manual. Hermeneutic analysis evidenced quantitative and qualitative aspects of the change, pointing also the problems in assessing change in adolescence.

The judges concluded that this is a clearly good outcome case, with a considerably to substantial degree of change, and which was substantially to completely due to the therapy. These conclusions provide further support to the effectiveness of the manualised TA treatment for depression in adults, and the case provides the first evidence that the manualised treatment was effective in treatment with a female adolescent with comorbid depression and anxiety.

This provides a new line of evidence for the effectiveness of the manualised TA therapy, which has now demonstrated effectiveness with adults and adolescents, with both 'straightforward' depression as well as comorbid anxiety, in two European cultures. The case also provides some initial indications that TA therapy may be effective for self-injuring behaviours and substance use problems amongst adolescents and further research is warranted to investigate this further.

Creating an early therapeutic alliance, supporting self-esteem, changing self-critical internal dialogues, developing an internal Nurturing Parent, providing appropriate permission tailored to the specific needs of the client and developing problem-solving ability all appeared to be mediators of change in this case, which were moderated by the intelligence of the client, her introspective capacity and the relational networks. These mediator and moderator factors all suggest areas for further research and can be used to examine and aggregate findings from all HSCED investigations of manualised TA therapy conducted so far.

## 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 in adulthood (Widdowson, 2016) has been effective in treating a major depressive disorder in an adolescent treated by a female therapist. 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 the manualised TA psychotherapy for depression as applied to adolescents.

## 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](mailto:enrico.benelli@unipd.it)

**Maddalena Bergamaschi**, Provisional Teaching and Supervising Transactional Analyst (Psychotherapy)

**Cristina Capoferri**, Certified Transactional Analyst Trainer (Psychotherapy)

**Stefano Morena**, Teaching and Supervising Transactional Analyst (Psychotherapy), President of ITACA - International Transactional Analysts for Childhood and Adolescence

**Vincenzo Calvo**, Assistant Professor, University of Padua

**Stefania Mannarini**, Associate Professor, University of Padua

**Arianna Palmieri**, Assistant Professor, University of Padua

**Mariavittoria Zanchetta**, Psychologist, University of Padua

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

## References

- Achenbach, T.M. (1991). *Manual for the Youth Self Report form and 1991 profile*. Burlington: Department of Psychiatry, University of Vermont.
- Achenbach, T. M., & Rescorla, L. (2001). *ASEBA school-age forms & profiles*.
- 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
- 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.
- Berne, E. (1961). *Transactional analysis in psychotherapy: a systematic individual and social psychiatry*. Grove Press. New York.
- Berne, E. (1972) *What do you say after you say hallo?* 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.
- 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.
- Costello, Jane E., Erkanli, A., & Angold, A. (2006). Is there an epidemic of child or adolescent depression?. *Journal of Child Psychology and Psychiatry*, 47(12), 1263-1271.
- 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., 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
- 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.
- Kapur, R. (1987). Depression: an integration of TA and psychodynamic concepts. *Transactional Analysis Journal*, 17:29-34.
- Kerr, C. (2013). TA Treatment of Emetophobia – A Systematic Case Study – 'Peter'. *International Journal of Transactional Analysis Research*, 4:2, 16-26.
- 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.
- McNeel, J. R. (2010). Understanding the power of injunctive messages and how they are resolved in rededuction therapy. *Transactional Analysis Journal*, 40(2), 159-169.
- Murray, C. J., Vos, T., Lozano, R., Naghavi, M., Flaxman, A. D., Michaud, C., ... & Aboyans, V. (2013). Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, 380(9859), 2197-2223.
- 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.
- Seligman, L. D., & Ollendick, T. H. (1998). Comorbidity of anxiety and depression in children and adolescents: An integrative review. *Clinical Child and Family Psychology Review*, 1(2), 125-144.
- 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
- Stewart, I., & Joines, V. (1987). *TA today: A new introduction to transactional analysis*. Vann Joines.
- Stewart, I., & Joines, V. (2012). *TA today: A new introduction to transactional analysis* (2nd edn). Lifespace. Nottingham.
- Stiles, W. B., Hill, C. E., & Elliott, R. (2015). Looking both ways. *Psychotherapy Research*, 25(3), 282-293.
- Van Noorden, M. S., Van Fenema, E. M., Van der Wee, N. J., Zitman, F. G., & Giltay, E. J. (2012). Predicting outcome of depression using the depressive symptom profile: the Leiden Routine Outcome Monitoring Study. *Depression and anxiety*, 29(6), 523-530.
- 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
- Wesselhoeft, R., Pedersen, C. B., Mortensen, P. B., Mors, O., & Bilenberg, N. (2015). Gender–age interaction in incidence rates of childhood emotional disorders. *Psychological medicine*, 45(04), 829-839.
- 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.