J. Behav. Ther. & Exp. Psychiat. 43 (2012) 770–779

Contents lists available at SciVerse ScienceDirect

Journal of Behavior Therapy and Experimental Psychiatry

journal homepage: www.elsevier.com/locate/jbtep

Think Effectively About Mood Swings (TEAMS): A case series of cognitive–behavioural therapy for bipolar disorders

Ruth Searson [a,1], Warren Mansell [a,*], Ian Lowens [b], Sara Tai [a]

[a] School of Psychological Sciences, Coupland I, University of Manchester, Oxford Road, Manchester M13 9PL, UK
[b] Salford Early Intervention in Psychosis Service, Greater Manchester West Mental Health NHS Trust, Pendleton House, Broughton Road, Salford, Manchester M6 6LQ, UK

ARTICLE INFO

Article history:
Received 18 March 2011
Received in revised form
16 June 2011
Accepted 6 October 2011

Keywords:
Bipolar disorder
Cognitive–behavioural therapy
Case series
Integrative model

ABSTRACT

Background and objectives: An integrative cognitive model for bipolar disorder proposes that multiple and extreme appraisals of changes in internal state and their reciprocal impact on behaviour, physiology and the environment provide the core mechanism in maintaining and escalating bipolar symptoms (Mansell, Morrison, Reid, Lowens, & Tai, 2007a).
Methods: A case series of cognitive–behavioural therapy (CBT) based on this model, known as the TEAMS approach (Think Effectively About Mood Swings), with seven participants was conducted. An A–B direct replication design with multiple baseline and follow-up assessments at one, three and six months was used. Treatment involved 12 sessions of CBT with an emphasis on addressing extreme positive and negative appraisals of internal state change.
Results: Improvements were reported for symptoms, functioning, cognitions and self-critical processes with large effect sizes on a range of measures, especially depression, at end of therapy and one-month follow-up. Five participants also showed clinically significant change in depression at both time-points.
Conclusions: This study provides preliminary evidence for the feasibility, acceptability and efficacy of CBT based on this model which warrants further evaluation.

1. Introduction

In comparison to cognitive models that have informed therapy for other psychological problems (such as anxiety disorders, trauma and psychosis), the progress of CBT for bipolar disorder has lagged by around 10–15 years. Existing CBT approaches for bipolar disorder predominantly focus on a combination of relapse prevention and traditional cognitive–behavioural strategies based on biopsychosocial models, delivered within a manualised structure (e.g. Lam, Jones, Hayward, & Bright, 1999). CBT trials based on these approaches have been effective in reducing bipolar symptoms, improving social functioning and reducing rates of relapse post-therapy (e.g. Ball et al., 2006; Lam et al., 2003). As such, this approach has been recommended by the National Institute for Health and Clinical Excellence (NICE, 2006).

However, the positive effects (including for depressive symptoms and relapse rates) have not always been maintained at long-term follow-up (Ball et al., 2006; Lam, Hayward, Watkins, Wright, & Sham, 2005). Furthermore, a large trial by Scott et al. (2006), which used a clinically heterogeneous sample, found that CBT was effective (through post-hoc analysis), but only for individuals with fewer than 12 episodes. As concluded by Jones (2004), further psychotherapy research needs to be based on clearer theoretical models of bipolar disorder, and needs to clarify which aspects of therapy are appropriate for which phases of the disorder, which a model-driven therapy may be more able to explore.

Generally, CBT is designed to work on current collaboratively identified problems. Treatment is focused on developing a formulation of how thinking styles and behaviours maintain and escalate current symptoms within unipolar depression (e.g. Watkins et al., 2007), anxiety disorders (e.g. Ehlers & Clark, 2000), and eating disorders (Fairburn, Cooper, & Shafran, 2003). Critically, bipolar patients report that their recovery involves not simply remaining free of relapse, but also regaining a sense of purpose in their lives and facing longstanding problems (Mansell, Powell, Pedley, Thomas, & Jones, 2010). Taking these principles on board, the aims of the integrative cognitive model for mood swings and bipolar disorders (Mansell, Morrison, Reid, Lowens, & Tai, 2007a) are to develop understanding of key psychological processes involved at different stages across bipolar spectrum disorders. In

* Corresponding author. Tel: +44 (0) 161 275 8589.
E-mail address: warren.mansell@manchester.ac.uk (W. Mansell).
URL: http://www.psych-sci.manchester.ac.uk/staff/131406
[1] Present address: Bolton Early Intervention in Psychosis Service, Paragon House, Chorley New Road, Bolton BL6 6HG, UK.

0005-7916/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jbtep.2011.10.001

keeping with advances in other fields, such as psychosis (e.g. Morrison, 2001), this model provides an accessible, formulation-based approach.

The model and its clinical application have been described in detail elsewhere (Mansell, 2007, 2010; Mansell & Hodson, 2009; Mansell, Morrison, Reid, Lowens, & Tai, 2007b). It has many influences, including first to third wave CBT and also control theory (Carver & Scheier, 1990; Powers, 1973). To summarise briefly, the model proposes that multiple and contradictory interpretations of internal state changes and their reciprocal impact on behaviour, physiology and the social environment are the core mechanisms involved in maintaining and escalating bipolar symptoms. Therefore, the defining feature of this model is that the individual experiences extreme positive *and* negative appraisals of the same internal states. For this reason, the intervention is named *Think Effectively About Mood Swings* (TEAMS) – with the aim of CBT to identify these interlinking but conflicting appraisals, thereby enabling more effective approaches to working with the complexity of internal state changes. For example, states of high activation could receive both positive and catastrophic appraisals such as "my fast thinking means I will solve all my problems" *and* "my racing thoughts mean I am about to lose control of my mind."

Whichever appraisal is accessed at any one time will depend on the dynamic interaction between the particular internal state and the current context (Mansell, 2010). These appraisals prompt efforts to control or enhance the internal state through "ascent" behaviours (which increase activation) or "descent" behaviours (which decrease activation). However, paradoxically, these behaviours contribute to the change in internal state, thus maintaining and heightening either manic or depressive symptoms respectively, and confirming dysfunctional appraisals. For example, the appraisal that "my fast thinking means I will solve all my problems" could lead to ascent behaviours such as increased activity, reduced sleep or ignoring others' advice to slow down, leading to a further activated state. On the other hand, the appraisal that "my racing thoughts mean I am about to lose control of my mind" could lead the individual to try and control their mood state through descent behaviours including worrying, ruminating, withdrawing and being self-critical, which could lead to a deactivated state and depressed mood. The model is represented diagrammatically in Fig. 1.

There is a substantial empirical evidence-base for the model. The Hypomanic and Positive Predictions Inventory (HAPPI; Mansell, 2006) was constructed to measure key appraisals of the model. The HAPPI has been shown to distinguish bipolar groups from non-clinical controls (Mansell, 2006; Mansell & Jones, 2006) and from people with remitted unipolar depression (Alatiq, Crane, Williams, & Goodwin, 2010; Mansell et al., 2011) and to prospectively predict symptoms of bipolar disorder over one month when controlling for clinical variables (Dodd, Mansell, Morrison, & Tai, 2011). The model is also consistent with experimental (Mansell & Lam, 2006) and qualitative (Mansell et al., 2010) studies of bipolar disorder.

The primary goal of therapy based on the model is to facilitate clients' awareness of their extreme attempts to exert control over their mood states and the appraisals that underlie these attempts. Further aims include helping clients develop more functional and adaptive ways of responding to changing internal states. The therapy follows a hierarchy described as the Pyramid of Principles (see Fig. 2). For example, it would be necessary for the therapist and client to reach a basic level of engagement before the client is likely to be able to talk about their experiences. However, it is anticipated that over the course of therapy, clients might move up and down the stages as strains and ruptures occur within the working alliance, and the client experiences therapeutic gains.

Key examples of techniques include: (1) finding a '*middle ground*' – drawing out a continuum of self-states from depressed to manic within which the client is encouraged to describe in detail states that lie between the two thereby promoting the pursuit and construction of 'healthy self-states' whose pros and cons can be evaluated to help them to achieve their personal goals; (2) alternative responses to ascent and descent behaviours may include noticing and accepting a change in internal state and 'mindfully' choosing not to act in ways that exacerbate it; (3) metacognitive monitoring of processes such as worry, rumination and self-attacking (conceptualised as ascent or descent behaviours); (4) an array of CBT techniques such as behavioural experiments, pie charts, memory, metaphor and imagery restructuring (Mansell, 2010).

The clinical application of the model has been reported in published case studies (Mansell, 2007; Mansell et al., 2007b),

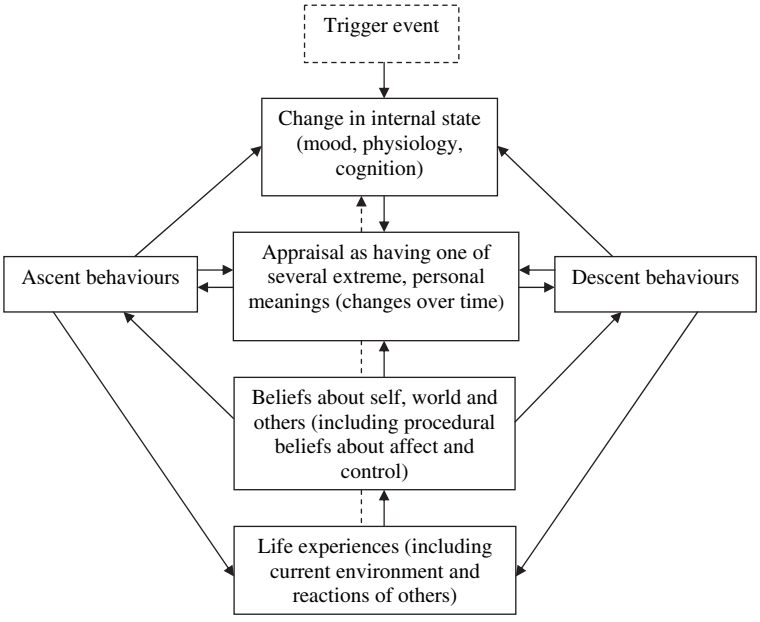

Fig. 1. The integrative cognitive model of mood swings and bipolar disorders.

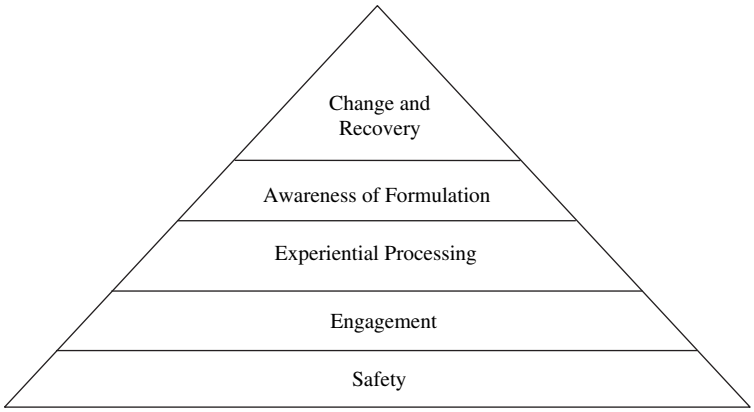

Fig. 2. The pyramid of principles.

which reported improvements to bipolar symptoms and key appraisals, thereby supporting the acceptability of the model. The next stage in building evidence for treatment based on the model was to formally evaluate it within a case series (Rice, 2008). The direct replication of a single case has the advantage of providing more complex data on individual changes when compared to a controlled study, which can only indicate the overall statistical change of a significant or non-significant result. Furthermore a case series can generalise treatment effects to similar individual clients (Barlow & Hersen, 1984). The case series can also identify more clearly the specific elements of therapy that impact on key symptoms and functioning by viewing the changes to individual scores through a session-by-session analysis.

The primary aim of the current study was to estimate the efficacy of a relatively brief (12 session) cognitive–behavioural therapy developed from the model for key bipolar symptoms, cognitions and general functioning. The objective was to assess whether clinically significant change could be achieved within 12 sessions, as used in a case series of CBT for residual depression (Watkins et al., 2007), and previous trials of CBT for other psychological disorders, including panic (Clark et al., 1994). However, it is acknowledged that the majority of previous CBT trials for bipolar disorder have utilised approximately 20 weekly sessions (Ball et al., 2006; Lam et al., 2003; Scott et al., 2006). The second aim of the study was to evaluate the feasibility and acceptability of this approach, as indicated by the number of participants able to complete the study, and by an analysis of participants' evaluations of therapy. Thirdly, the study aimed to explore changes to key psychological processes identified within the model; in particular, self-criticism. The main focus of this article is to report the clinical outcomes. Changes to other psychological processes are included.

2. Method

2.1. Participants

Seven participants (five females and two males) with a bipolar disorder diagnosis, as assessed by the Structured Clinical Interview for DSM-IV-TR Axis 1 Disorders (Research Version) (SCID-I; First, Gibbon, Spitzer, & Williams, 2002) were recruited to the case series. Exclusion criteria were: a diagnosis of schizophrenia; schizoaffective disorder; primary substance misuse (mood episodes caused purely by substance misuse); psychosis outside of mood episodes or a current episode of mania. All participants were recruited from Community Mental Health Teams (CMHTs) within the National Health Service (NHS) Trust where they were receiving treatment as usual. Participant ages ranged from 23 to 44 years. All participants were in receipt of medication: five were being treated with a mood stabiliser (and additionally, for four of these participants, an anti-depressant) and two received anti-psychotic medication. Further details of participants' history and initial presentation are included in Table 1. It should be noted that some clients showed a fluctuating pattern between low and high mood states later in therapy.

2.2. Design

The case series used an A–B direct replication design involving baseline, treatment and follow-up phases. An increasing multiple

Table 1
Participant details.

No.	Diagnosis	Approximate length of diagnosis	Brief history	Clinical presentation at screening (on SCID-I)
1	BP I	6 weeks	Manic episode (with psychotic features) and hospital admission 6 weeks earlier; 15 year history of reported low moods but no prior service input	MDE
2	BP II	17 months	Depressed four years ago following life stressors; later diagnosed with BP	MDE
3	BP II	5 weeks	Two episodes of hypomania followed by depression within last 18 months	MDE
4	BP II	12 years	16 year history of mood instability; one brief informal admission	Euthymic
5	BP II	2 years	12 year history of mood instability; several hospital admissions for low mood	MDE
6	BP I	6 years	Three hospital admissions for manic and hypomanic episodes; at least four past episodes of MDE	Euthymic
7	BP II	8 years	10 year history of mood fluctuations including mixed states; one previous admission	Hypomania (within last month); reported low mood by start of therapy

Note. BP I = Bipolar Disorder I; BP II = Bipolar Disorder II; MDE = Major Depressive Episode.

baseline strategy was employed to strengthen the design, so that participants attended between three and five weekly visits to complete questionnaires before commencing therapy. After the therapy phase, participants completed three follow-up visits at one, three and six months.

2.3. Measures and assessments

Participants completed a range of measures throughout the study to assess symptoms, cognitions, functioning and self-critical processes. Some of the briefer questionnaires were completed weekly throughout baseline, therapy and follow-up (ISS, PANAS, Client-HAPPI, WSAS and all measures of self-critical processes). Other measures were completed at six time-points only: at baseline visit 1, last baseline visit, end of therapy (session 12), and at one, three and six months' follow-up (BDI, BAI, Bech–Rafaelsen interview, CORE). The primary measures that were used as measures of clinical change were the ISS, BDI, HAPPI, CORE and WSAS.

2.3.1. Structured Clinical Interview for DSM-IV-TR Axis 1 Disorders (Research Version) (SCID-I; First et al., 2002)

The therapist administered the SCID following training through observation of training videos, role-play and feedback from a trained supervisor. Mood modules (A and D) of the SCID-I and screen for psychosis symptoms (B) were administered at screening to establish that participants had a diagnosis within the bipolar spectrum and did not meet exclusion criteria. SCID-I interviews were recorded with participants' consent and checked for reliability by the second author, which showed 100% agreement in relation to past and current mood episodes and diagnoses of bipolar I or II disorder. The SCID-I was also repeated at the three month follow-up visit.

2.3.2. Internal State Scale (ISS; Bauer et al., 1991)

This 16-item self-report scale assesses a range of bipolar symptoms over the previous 24 h, with each item rated from 0 to 100 along a visual analogue scale. For this study, a composite score was computed comprised of the Activation, Depression and Conflict subscale totals.

2.3.3. Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961)

This 21-item scale measures current severity of depression over the past week. Each item is scored from 0 to 3 with a maximum score of 63.

2.3.4. Beck Anxiety Inventory (BAI; Beck, Epstein, Brown, & Steer, 1988)

This 21-item scale measures the severity of anxiety over the past week. The items include physical and cognitive symptoms and are rated from 0 to 3 with a maximum score of 63.

2.3.5. Bech–Rafaelsen semi-structured interview (combining Bech–Rafaelsen Mania Rating Scale (MRS; Bech, Rafaelsen, Kramp, & Bolwig, 1978) and the Hamilton Depression Rating Scale (HDRS; Hamilton, 1960))

The interview involves questions about 17 depressive symptoms and 11 manic symptoms. Each item is rated through the use of questioning and observation from either 0 to 4 or 0 to 2 representing increasing levels of severity. All interviews were recorded and a sample was rated by a supervisor of the project (ST) to assess interrater reliability. Using the Kappa statistic, it was found that there was an 'outstanding' level of agreement for 5 interviews and a 'moderate' level for one interview (Landis & Koch, 1977).

2.3.6. Positive and Negative Affect Scale (PANAS; Watson, Clark, & Tellegen, 1988)

A 20-item scale of different emotions and feelings (10 positive and 10 negative) rated from 1 to 5 for occurrence over the past week.

2.3.7. Hypomanic and Positive Predictions Inventory (HAPPI; Dodd, Mansell, Sadhnani, Morrison, & Tai, 2010)

The HAPPI is a measure of the multiple and extreme appraisals of internal states central to the model and was used as the key measure to assess changes to these cognitions across the course of therapy. The 61-item version was administered at each baseline visit (to establish the cognitions most relevant for each client), and thereafter at end of therapy and at follow-up. Each item is rated from 0 to 100 to indicate level of conviction in each item from 0% (do not believe at all) to 100% (believe this completely).

2.3.8. Client-HAPPI

This shorter individualised version of the HAPPI (approximately 10–15 items) was developed for each client using the full HAPPI ratings at baseline to identify and then monitor the most important cognitions for each participant during the therapy phase.

2.3.9. Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM; Evans et al., 2000)

This 34-item scale monitors psychological distress across four domains: subjective well-being; specific problems; functioning and risk to self and others. Each item is scored from 0 to 4. Clinical scores are calculated as the mean of all items for each domain and for the total scale, which are then multiplied by 10 so that scores are expressed in whole numbers from 0 to 40. A cut-off score of 10 is recommended to distinguish a general population sample from a clinical sample (Connell et al., 2007), although it was recognised that this may be higher for some diagnoses.

2.3.10. Work and Social Adjustment Scale (WSAS; Marks, 1986)

This scale assesses perceived level of impairment in five areas: Work, Home management, Social life, Private leisure and Relationships. Each item is scored from 0 (no impairment) to 8 (very severe impairment), with a total maximum score of 40.

2.3.11. Self-Compassion Scale (SCS; Neff, 2003)

This 26-item scale includes six subscales: Self-kindness; Self-judgement; Common humanity; Isolation; Mindfulness and Over-identification. Responses are scored from 1 to 5 to produce a mean score for each subscale and an overall self-compassion score (maximum of 30).

2.3.12. Forms of Self-Criticising/Attacking and Self-Reassuring scale (FSCRS; Gilbert, Clarke, Hempel, Miles, & Irons, 2004)

This 22-item scale includes two components of self-criticism: 'Inadequate-self' and 'Hated-self', and a 'Reassured-self' subscale. Items are scored from 0 to 4 to produce a total score for each subscale.

2.3.13. Functions of Self-Criticising/Attacking Scale (FSCS; Gilbert et al., 2004)

This 21-item scale comprises the 'Self-correction' and 'Self-persecution' components. Items are scored from 0 to 4 to produce a total score for each component.

2.3.14. California Psychotherapy Alliance Scale – patient version (CALPAS-P; Gaston & Marmar, 1993)

This 24-item scale assesses the nature of the therapeutic alliance which includes four subscales (six items each): Patient

Commitment (PC); Patient Working Capacity (PWC); Therapist Understanding and Involvement (TUI); and Working Strategy Consensus (WSC). Items are rated from 1 to 7 to produce subscale mean scores and a total score (from 4 to 28). The CALPAS-P was administered once at one-month follow-up to assess participant perceptions of the therapeutic relationship as part of assessing the acceptability of this style of CBT.

Additionally a qualitative feedback form was administered at one-month follow-up.

2.4. Procedure

Ethical approval was obtained from the Local Research Ethics Committee to conduct the study within an NHS Trust. The study also received approval from the University Ethics Committee and NHS Trust R&D approval. Following an initial screening appointment, participants were allocated to between 3 and 5 weeks' baseline. During baseline, participants attended weekly appointments of approximately 30 min (at their usual CMHT base) to complete assessments. The therapy phase consisted of 12 individual CBT sessions of approximately 50 mins' duration, which were based upon an individually formulated version of the model for each client. Participants attended post-therapy follow-up appointments at one, three and six months to complete further assessments (of between 30 and 60 mins' duration). Throughout the study, due to the large number of measures used and time limits, it was agreed that self-report questionnaires could be completed at home prior to the appointment. The delivery of the therapy and all assessments completed at screening, baseline and follow-up and all client contacts were provided by the first author. Therapy sessions were recorded (with participant consent) to monitor adherence to the model and for use in supervision. The first author received substantial training in use of the model and in delivering relevant assessments from all other authors prior to recruitment. Weekly clinical supervision was provided by the second author including detailed feedback on a selected therapy session. The first author regularly liaised with GPs and other relevant professionals within the client's CMHT on the progress of therapy and in the event of risk and followed Trust policies in the recording and sharing of clinical notes.

3. Results

3.1. Feasibility

All seven participants completed the full course of therapy and one-month follow-up visit. Five participants completed three and six month follow-up visits. Two participants were lost to follow-up at this stage due to non-engagement.

3.2. Clinical outcome

The distribution of mean baseline, end of therapy scores, one-month follow-up scores, baseline to end of therapy change scores and baseline to one-month follow-up change scores were not markedly skewed for the majority of the data. Therefore effect sizes were calculated for end of therapy and one-month follow-up. However nonparametric descriptions are included for scales showing moderate skewness (see Table 2). Effect sizes (Cohen's *d*) were calculated by dividing the mean change in individual scores (from baseline to end of therapy or one-month follow-up) by the pooled standard deviation (SD) of scores at these time-points. The pooled standard deviation is calculated as $\sqrt{[(SDpre^2 + SDpost^2)/2]}$, where 'pre' refers to mean baseline scores and 'post' to end of

Table 2
Changes to sample mean scores. Bold values highlights the effect sizes.

Measure	Baseline 1 ($N=7$) *M* (SD)	Last baseline ($N=7$) *M* (SD)	Mean baseline ($N=7$) *M* (SD)	End of therapy ($N=7$) *M* (SD)	Effect size pre–post *d*	1 month follow-up ($N=7$) *M* (SD)	Effect size pre-1m FU *d*	3 month follow-up ($N=5$) *M* (SD)	6 month follow-up ($N=5$) *M* (SD)
BDI	24.57 (6.50)	18.29 (7.95)	21.43 (4.47)	9.00 (6.00)	**2.35**	6.43 (5.62)	**2.95**	11.00 (9.46)	5.80 (5.40)
BAI	13.29 (8.90)	14.00 (6.08)	13.64 (3.46)	13.71 (11.91)	**0.01**	11.71 (10.34)	**0.25**	16.20 (16.81)	12.00 (10.39)
BR (dep)	12.29 (6.63)	8.86 (6.62)	10.57 (4.52)	6.14 (5.11)	**0.92**	–	–	7.20 (3.77)	–
BR (man)[a]	3.00 (1.00, 4.00)	1.00 (0.00, 6.00)	2.00 (1.00, 5.00)	1.00 (0.00, 1.00)	–	–	–	1.00 (0.50, 1.50)	–
ISS			407.44 (153.95)	234.29 (241.43)	**0.86**	298.14 (175.10)	**0.66**	375.80 (160.70)	222.80 (231.76)
PANAS-P			22.37 (7.98)	23.86 (11.34)	**0.15**	28.57 (11.93)	**0.60**	22.60 (15.90)	30.80 (14.27)
PANAS-N[b]			24.53 (6.99)	17.57 (5.74)	**1.09**	17.86 (6.15)	**1.01**	20.40 (10.29)	16.20 (7.01)
HAPPI			49.96 (11.06)	28.29 (22.76)	**1.21**	25.57 (21.72)	**1.42**	24.20 (23.73)	20.40 (21.39)
Client-HAPPI[b]			78.60 (15.13)	44 (24.44)	**1.70**	44.29 (27.87)	**1.53**	39.2 (29.76)	33.8 (30.23)
CORE[b]	18.71 (6.42)	14.71 (5.77)	16.71 (3.26)	8.71 (5.22)	**1.84**	10.03 (6.91)	**1.24**	12.50 (8.31)	7.00 (5.75)
WSAS			18.93 (5.94)	8.29 (7.68)	**1.55**	8.43 (8.73)	**1.40**	12.80 (11.08)	5.80 (3.19)
FSCRS-I			25.30 (6.86)	14.71 (10.77)	**1.17**	10.71 (7.83)	**2.90**	11.20 (10.04)	10.00 (8.22)
FSCRS-H			7.64 (3.54)	3.43 (3.00)	**1.27**	3.00 (2.77)	**1.46**	3.20 (3.35)	2.20 (3.27)
FSCRS-R			13.30 (4.44)	18.14 (4.85)	**1.04**	20.14 (7.38)	**1.12**	20.80 (10.09)	21.20 (7.19)
FSCS-SC			24.82 (11.01)	21.29 (14.98)	**0.27**	18.29 (12.13)	**0.56**	20.40 (15.50)	–
FSCS-SP[b]			8.20 (6.00)	5.14 (6.69)	**0.48**	3.29 (5.71)	**0.84**	8.20 (8.64)	–
SCS			12.40 (2.77)	18.43 (5.29)	**1.43**	19.86 (5.64)	**1.68**	19.60 (6.80)	20 (4.06)

M = mean; SD = standard deviation; BDI = Beck Depression Inventory; BAI = Beck Anxiety Inventory; BR (dep) and (man) = Bech–Rafaelsen depression and mania scales; ISS = Internal States Scale; PANAS-P and PANAS-N = Positive and Negative Affect Scale – Positive and Negative subscales; CORE = Clinical Outcomes in Routine Evaluation; WSAS = Work and Social Adjustment Scale; HAPPI = Hypomanic and Positive Predictions Inventory, FSCRS (I, H, R) = Forms of Self-Criticising/Attacking and Self-Reassuring Scale – Inadequate-self, Hated-self and Reassured-self subscales; FSCS (SC, SP) = Functions of Self-Criticising/Attacking Scale – Self-correction and Self-persecution subscales; SCS = Self-Compassion Scale.

Note: Mean baseline for each measure = mean of all participants' mean baseline scores. For BDI, BAI, CORE and Bech–Rafaelsen scales, mean baseline for each participant is the mean of two baseline visits (first and last); for weekly measures, mean baseline for each participant is calculated from all baseline visits (between 3 and 5).

[a] Median and interquartile range (IQR) reported only.

[b] Medians and IQR included for measures showing moderate skewness where effect sizes calculated: PANAS-N = 21.60 (20.67, 27.00) at mean baseline, 17.00 (13.00, 24.00) at end of therapy, 18.00 (11.00, 23.00) at 1 month follow-up; CORE = 17.50 (15.50, 19.50) at mean baseline, 8.00 (6.00, 10.00) at end of therapy, 8.50 (7.00, 14.00) at 1 month follow-up; Client-HAPPI = 83.67 (69.00, 88.75) at mean baseline, 53.00 (22.00, 68.00) at end of therapy, 55.00 (16.00, 63.00) at 1 month follow-up; FSCS-SP = 7.25 (3.80, 14.33) at mean baseline, 2.00 (0.00, 12.00) at end of therapy, 1.00 (0.00, 3.00) at 1 month follow-up.

therapy or follow-up scores. This procedure was set out by Cohen (1977) and has been used and described in other case series (Watkins et al., 2007; Wells & Sembi, 2004). Cohen (1977) identified effect sizes as small ($d = 0.2$), medium ($d = 0.5$) and large (0.8). Mean scores (and standard deviations) were calculated at three and six month follow-up where data for $n = 5$ are available. The data was not markedly skewed at these time-points but effect sizes were not calculated due to the smaller sample size. Graphs were constructed for the weekly outcome measures to demonstrate detailed changes across the whole course of therapy.

3.3. Symptoms

As shown in Table 2, the largest effects were for depression, as shown by large effect sizes on the BDI at both end of therapy and one-month follow-up, and on the Bech–Rafaelsen Depression Scale at end of therapy. There was also a large effect for a reduction in scores for negative feelings on the PANAS at both time-points. In relation to bipolar symptoms, there was a large effect on the ISS at end of therapy. There was no change on the BAI at end of therapy. On the SCID-I, at three months' follow-up, two participants (1 and 2) reported an increase in depressed mood of approximately three weeks' duration however neither participant met full criteria for a major depressive episode. Participant 7 did meet criteria for a major depressive episode within these three months, and participant 4 met criteria for a past hypomanic episode of approximately two months' duration. Participant 3 had not experienced any mood episodes within the three months since the end of therapy.

3.4. Key cognitions

Effect sizes on the HAPPI and Client-HAPPI were large at both end of therapy and one-month follow-up. As shown in Fig. 3, the scores for three participants (1, 2 and 7) reflect a fairly significant decreasing trend throughout the therapy phase which was

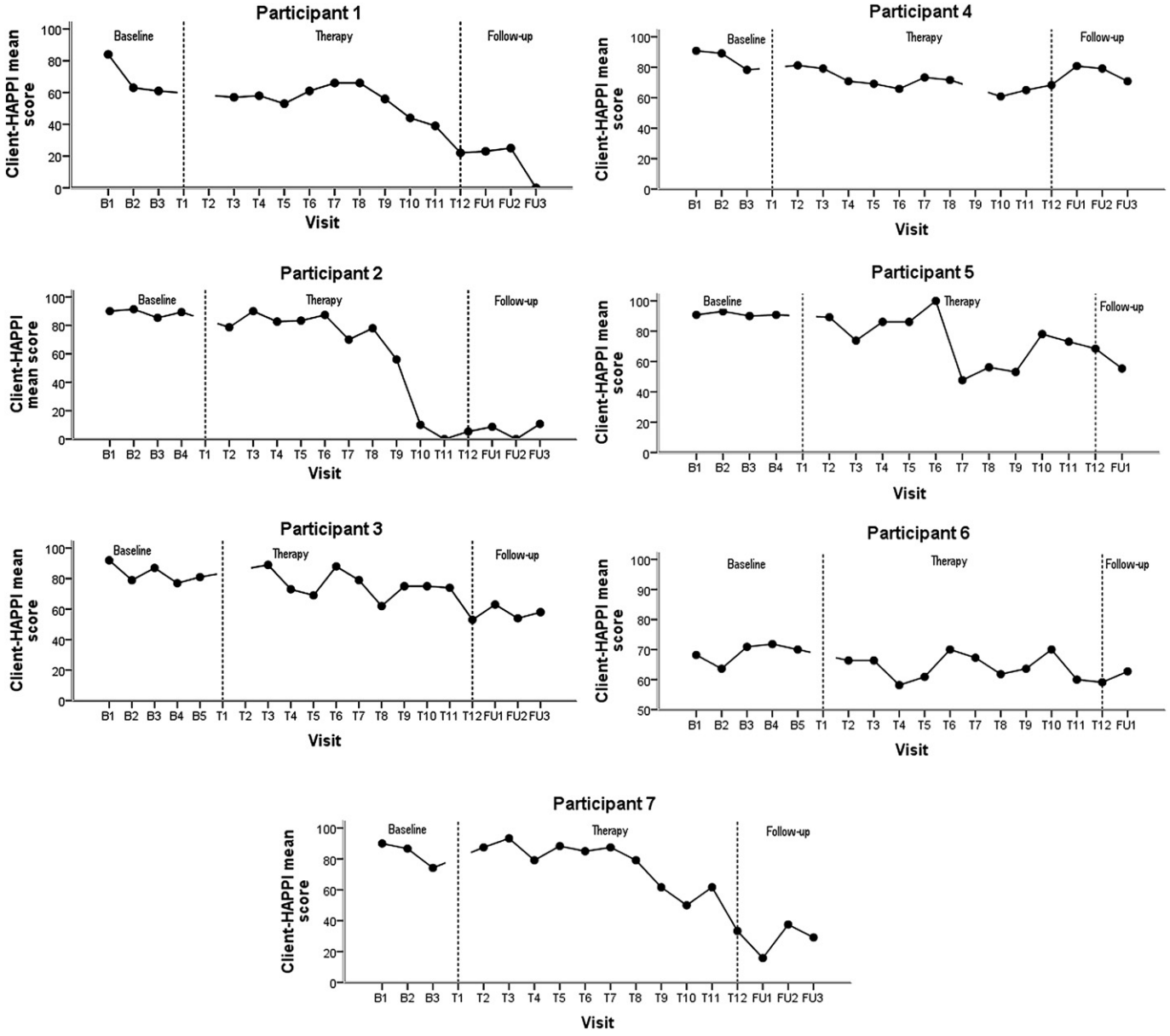

Fig. 3. Changes to participant scores on the Client-HAPPI.

maintained throughout the six month follow-up period. These graphs also largely show a pattern of baseline stability thus lending support to the reliability of change (Leslie & O'Reilly, 1999).

3.5. Functioning

There were large effect sizes for an improvement in general functioning on the CORE and WSAS at end of therapy and one-month follow-up. The total mean score on the CORE at end of therapy was also below the identified cut-off score of 10 between a clinical and general population (Connell et al., 2007) from above this score at baseline. As shown in Fig. 4, the individual graphs showing changes on the WSAS indicate that for five participants (1, 2, 3, 5 and 6), some scores within the baseline period had been above an identified clinical mean of 25 (Mundt, Clarke, Burroughs, Brenneman, & Griest, 2001), and for five participants (1, 2, 3, 5 and 7), scores at both end of therapy and one-month follow-up were below the cut-off score of 10, specified by Mundt, Marks, Shear, and Griest (2002).

3.6. Self-critical processes

As shown in Table 2, there were large effect sizes at both end of therapy and one-month follow-up on all subscales of the FSCRS and the SCS, and moderate effect sizes on the subscales of the FSCS. This reflects promising results that appear to be largely maintained over the follow-up period.

3.7. Clinically significant change

A clinical significance analysis was also performed, using the two-fold criterion *c* set out by Jacobson, Follette, and Revenstorf (1984).[2] This criterion requires the individual's score at post-treatment to move from outside the range of a clinical group to within the range of a 'functional' group by crossing a calculated 'cut-off point' and to demonstrate statistically reliable change. The cut-off point is calculated using normative data from the current study (clinical sample) and a non-clinical sample using the calculation $c = S_0M_1 + S_1M_0/S_0 + S_1$: where S_0 = standard deviation from the non-clinical group; M_1 = baseline mean from the current sample; S_1 = baseline standard deviation from the current sample; M_0 = mean from the non-clinical group. The non-clinical data was drawn from Beck (1967) for the BDI (mean = 10.9, SD = 8.1) and from Udachina and Mansell (2007) for the ISS (mean = 200.9, SD = 57.96). A reliable change index is also calculated for each measure.

Using this analysis, to be 'recovered', participants would need to cross the cut-off point and make reliable change in the direction of functionality. To be 'improved', participants have made reliable change in the direction of functionality but without crossing the cut-off point. Participants who show 'no change' have not made reliable change (regardless of whether they cross the cut-off point) and to have 'deteriorated' participants have made reliable change but in the opposite direction of functionality.

On the BDI, five participants (1, 2, 3, 5 and 6) were classed as clinically 'recovered' at end of therapy and one-month follow-up. A further participant (7) was 'improved' at both time-points and one participant (4) showed no change at either time-point. On the ISS, two participants (2 and 7) were 'recovered' at both end of therapy and one-month follow-up. Two participants (5 and 6) were 'improved' at end of therapy however three participants (1, 3 and 4) showed no change and by one-month follow-up five participants (1, 3, 4, 5, 6) showed no change. On the WSAS, three participants (4, 6, 7) showed no change at both end of therapy and one-month follow-up. Two participants were 'improved' (1, 5) and two were 'recovered' (2, 3) and by one-month follow-up all four of these participants were 'recovered.'

3.8. Acceptability

On the CALPAS-P subscales, the mean score for the group (from a possible range of 1–7) was 5.45 (SD = 0.72) for Patient Commitment, 5.98 (SD = 0.66) for Patient Working Capacity, 6.43 (SD = 0.74) for Therapist Understanding and Involvement and 6.36 (SD = 0.64) for Working Strategy Consensus. The total mean score for the group (from a possible range of 4–28) was 24.21 (SD = 1.8) indicating a high level of therapeutic alliance. On the qualitative feedback form, all seven participants indicated that the therapy had been helpful and provided positive feedback on their experience of therapy (with mixed feedback from one participant). Mean ratings on how helpful different aspects of therapy had been ranged from 3.7 to 4.6 (from a possible range of 1–5) with ratings of 4.6 for 'formulation' and 'relationship with therapist.' All participants engaged with the whole course of therapy suggesting that this style of CBT was acceptable to clients.

4. Discussion

The case series design has a limited capacity to demonstrate efficacy of a treatment as it does not involve a control group. However, with the utilisation of multiple baseline, session-by-session measures and the reporting of effect sizes, judgements can be made about the potential impact of an intervention to justify further, controlled evaluation.

The large effect sizes at end of therapy and one-month follow-up suggest that there were improvements to symptoms (particularly depression), key cognitions and psychosocial functioning. As a point of comparison, the baseline to post-treatment effect size for the BDI ($d = 2.35$) compares favourably to the effect size of a successful trial that also assessed depression using the BDI (CBT $d = 0.84$; TAU $d = 0.47$; Ball et al., 2006). There were also large improvements on measures of self-attacking and self-compassion at end of therapy and one-month follow-up.

From viewing changes to mean scores in Table 2, there appears to be a pattern of further improvement or reasonable stability at one-month follow-up in comparison to end of therapy. However, for measures of symptoms and functioning, there appears to be a trend of worsening scores at three months' follow-up, followed by improvement at six months' follow-up (to be either consistent with or further improved than the end of therapy). Alternatively, the measures of cognitive processes (including the HAPPI, Client-HAPPI, Reassured-self subscale of the FSCRS and SCS) show a pattern of continued or further improvement across the whole follow-up period. It is of note that there were no changes on the BAI. However, when recruited to the study, all seven participants reported feeling depressed in mood and highlighted this as a priority for therapy. Four participants met full criteria for major depressive episode on the SCID-I at screening. Therefore, the formulations that were developed for all participants initially focused on low mood cycles. As a result although formulations of 'high' mood were developed for all participants, there was less opportunity within 12 sessions to thoroughly target intervention on activated mood states that did not present as frequently during the course of therapy. This observation may relate to the limited changes on the BAI in that 'high' mood states can include physical symptoms similar to those experienced in anxiety and ratings on

[2] Criterion *c* was used for the BDI and ISS. For the WSAS the more stringent criterion *a* was used (where the cut-off point is defined as the clinical mean ± 2 SD in the direction of functionality) as non-clinical normative data was not available.

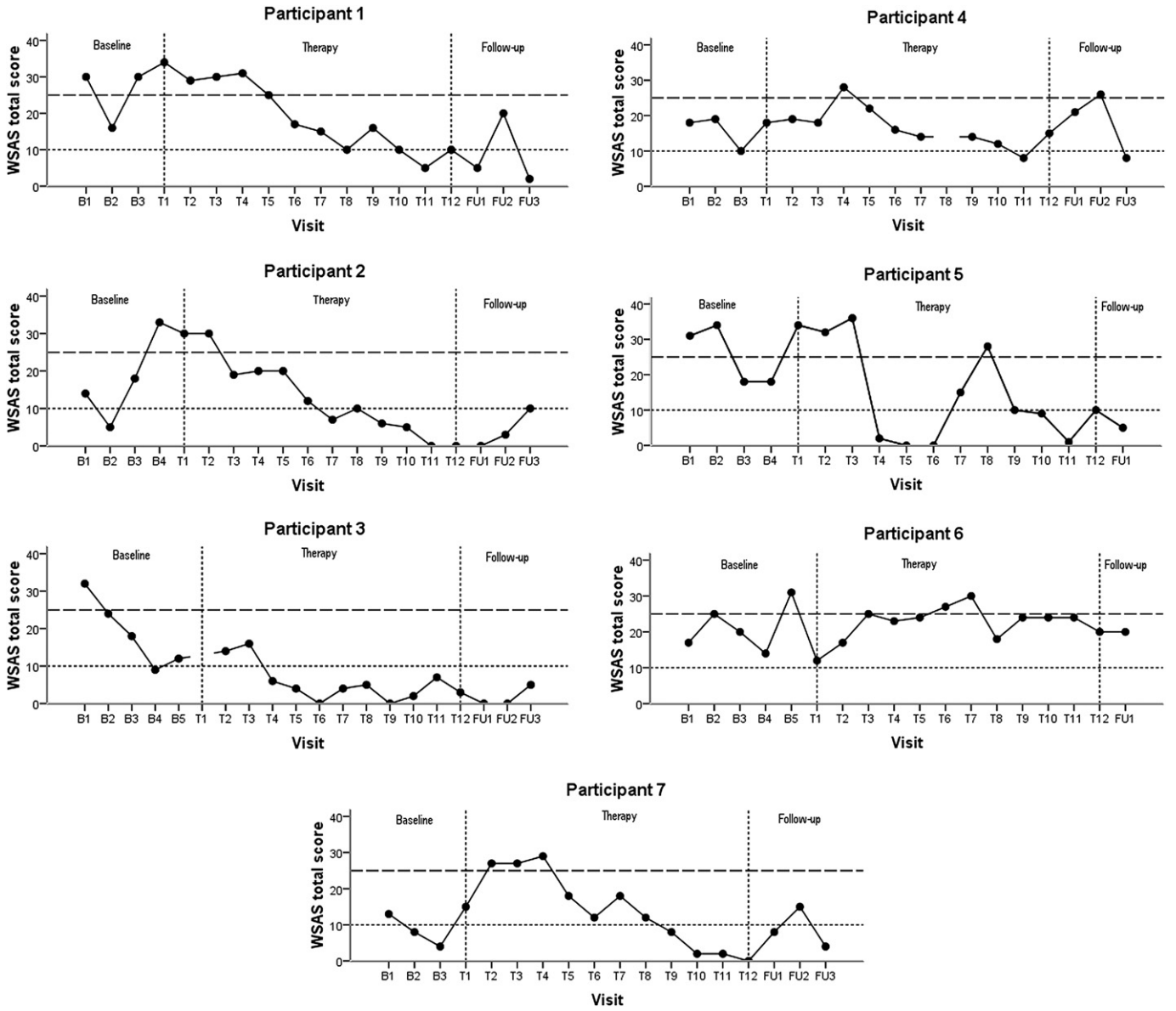

Fig. 4. Changes to participant scores on the Work and Social Adjustment Scale (WSAS). Note: Horizontal lines indicate a clinical mean and cut-off point from validated and published data. The clinical mean (dashed line) of 25 (SD = 8.23) is drawn from a depression group (receiving treatment as usual) at baseline ($n = 124$; Mundt et al., 2001). A cut-off score of 10 (dotted line) between a clinical and general population was identified by Mundt et al. (2002) from the data of two studies (total $n = 577$; Griest et al., 2002; Mundt et al., 2001).

the BAI may have reflected agitation in relation to high mood. There are also high levels of comorbidity of anxiety disorders in bipolar disorder (Simon et al., 2004). It is also acknowledged that anxiety was not specifically targeted in therapy with the majority of participants. Furthermore it may be that the BAI is more a measure of physical anxiety, and as such, it may have been beneficial to monitor worry about mood states which may have produced different effects.

At the three month follow-up, three out of five participants had not experienced any mood episodes (as assessed by the SCID-I) since the end of therapy. Individually, five of the seven participants were also 'clinically recovered' on the BDI at end of therapy and one-month follow-up, and four participants were 'recovered' on the WSAS at one-month follow-up. As proposed by Jacobson et al. (1984), reliance on group means and statistical significance tests only in evaluating treatment effects does not highlight both the variability of individual changes and whether changes are clinically meaningful. The total mean score for the sample on the CORE at end of therapy was also below the identified cut-off score of 10 between a clinical and general population (Connell et al., 2007) from above this score at baseline. These findings highlight that it is particularly important in both research and clinical practise to evaluate functioning as an outcome measure and not just symptom change due to the large impact on social functioning that exists in bipolar disorder. Therefore working with clients to reclaim functional outcomes should remain an important target for therapy. The improvements in functioning may also relate to the large effects that were found for depression. Considering also the improvements to self-critical processes, this suggests that this therapy may also be an effective intervention for residual depression.

There are several limitations associated with a case series design including the lack of a control group and the influence of

confounding variables that could not be controlled. This results in difficulties in establishing conclusions about the effectiveness of therapy as the 'active ingredient' over and above the non-specific effects (such as spending time with a therapist) and the potential occurrence of external factors. However, the use of a multiple baseline can negate these effects to a degree by building non-specific factors into the baseline period, such as therapist contact.

The majority of participants were on medication and it was deemed unethical to request the stabilisation of medication throughout the study. Participants were also receiving input from other professionals. Life events could have also impacted on symptoms. Furthermore, it is recognised that there may be significant differences in terms of participants' mental health experiences, including specific diagnosis, number of previous episodes and whether some participants had previously received CBT. However none of the participants had received input using the current model.

This study had a large number of outcome measures. Due to the small sample size, it is acknowledged that statistically this can increase the chances of error. The small sample size also limits the degree to which the results are generalisable to a 'typical' population. However this sample included a range of ages (23–44 years), gender (5 females and 2 males), diagnoses within the bipolar spectrum and participants from different social, cultural, ethnic and educational backgrounds, suggesting that the sample may have been representative of a clinical population.

The limitations of using self-report measures and the fact that both therapist and supervisor ratings on the interview measures were not blind to the aims of the study, or to the specific time-points at which these assessments were being completed is also acknowledged. Furthermore, it is recognised that the therapist had a dual role in also being an assessor in this study.

5. Conclusions

The results of this case series show improvements to symptoms (notably depression), psychosocial functioning, key cognitions and self-critical processes. The majority of the sample also showed 'clinically significant' improvement for depression at end of therapy and one-month follow-up, and for functioning at one-month follow-up. The therapy based on the model was found to be feasible and acceptable as evidenced by high retention rates in therapy and follow-up and by participant feedback.

These encouraging findings suggest the need to further test out the model within a larger controlled study to more accurately conclude the impact of the therapy. It is recommended that the next step would be a pilot trial before proceeding with a larger multi-site randomised controlled trial. As part of this research, it will be helpful to further explore and define the 'active ingredients' of the therapy. An important area for research would be to specifically explore how this therapy may be different to previous CBT approaches for bipolar disorder. Although theoretically this model is distinct from previous approaches and there are likely to be subtle differences in the style and focus of the therapy, it is acknowledged that there are also likely to be similarities with other forms of CBT for bipolar disorder.

Acknowledgments

This study was funded by the University of Manchester as part of the first author's doctorate thesis. Dr Warren Mansell is supported by the ESRC Emotion Regulation of Others and the Self Network Grant (RES 060-25-0044). We would also like to thank the participants for their contribution to this research and the participating NHS Trust for their support with recruitment.

References

Alatiq, Y., Crane, C., Williams, J. M. G., & Goodwin, G. M. (2010). Dysfunctional beliefs in bipolar disorder: hypomanic vs. depressive attitudes. *Journal of Affective Disorders, 122*, 294–300.

Ball, J. R., Mitchell, P. B., Corry, J. C., Skillecorn, A., Smith, M., & Malhi, G. S. (2006). A randomized controlled trial of cognitive therapy for bipolar disorder: focus on long-term change. *Journal of Clinical Psychiatry, 67*, 277–286.

Barlow, D. H., & Hersen, M. (1984). *Single case experimental designs: Strategies for studying behaviour change* (2nd ed.). Oxford: Pergamon Press.

Bauer, M. S., Crits-Christoph, P., Ball, W. A., Dewees, E., McAllister, T., Alahi, P., et al. (1991). Independent assessment of manic and depressive symptoms by self-rating. Scale characteristics and implications for the study of mania. *Archives of General Psychiatry, 48*, 807–812.

Bech, P., Rafaelsen, O. J., Kramp, P., & Bolwig, T. G. (1978). The Mania Rating Scale: scale construct and inter-observer agreement. *Neuropharmacology, 17*, 430–431.

Beck, A. T. (1967). *Depression: Causes and treatment*. Philadelphia: University of Pennsylvania Press.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: psychometric properties. *Journal of Consulting and Clinical Psychology, 56*, 893–897.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*, 561–571.

Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: a control-process view. *Psychological Review, 97*, 19–35.

Clark, D. M., Salkovskis, P. M., Hackmann, A., Middleton, H., Anastasiades, P., & Gelder, M. G. (1994). A comparison of cognitive therapy, applied relaxation and imipramine in the treatment of panic disorder. *British Journal of Psychiatry, 164*, 759–769.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (revised ed.). New York: Academic Press.

Connell, J., Barkham, M., Stiles, W. B., Twigg, E., Singleton, N., Evans, O., et al. (2007). Distribution of CORE-OM scores in a general population, clinical cut-off points and comparison with the CIS-R. *British Journal of Psychiatry, 190*, 69–74.

Dodd, A. L., Mansell, W., Morrison, A. P., & Tai, S. (2011). Extreme appraisals of internal states and bipolar symptoms: The Hypomanic Attitudes and Positive Predictions Inventory. *Psychological Assessment, 23*, 635–645.

Dodd, A. L., Mansell, W., Sadhnani, V., Morrison, A. P., & Tai, S. (2010). Principal components analysis of the Hypomanic Attitudes and Positive Predictions Inventory and associations with measures of personality, cognitive style and analogue symptoms in a student sample. *Behavioural and Cognitive Psychotherapy, 38*, 15–34.

Ehlers, A., & Clark, D. M. (2000). A cognitive model of posttraumatic stress disorder. *Behaviour Research and Therapy, 38*, 319–345.

Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., et al. (2000). CORE: Clinical Outcomes in Routine Evaluation. *Journal of Mental Health, 9*, 247–255.

Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: a "transdiagnostic" theory and treatment. *Behaviour Research and Therapy, 41*, 509–528.

First, M. B., Gibbon, M., Spitzer, R. L., & Williams, J. B. W. (2002). *Structured Clinical Interview for DSM-IV-TR axis 1 disorders (Research version)*. New York: Biometrics Research Department, New York State Psychiatric Institute.

Gaston, L., & Marmar, C. R. (1993). Manual of California Psychotherapy Alliance Scales (CALPAS). Unpublished manuscript.

Gilbert, P., Clarke, M., Hempel, S., Miles, J. N. V., & Irons, C. (2004). Criticizing and reassuring oneself: an exploration of forms, styles and reasons in female students. *British Journal of Clinical Psychology, 43*, 31–50.

Griest, J., Marks, I., Baer, L., Kobak, K., Wenzel, K., Hirsch, M., et al. (2002). Behaviour therapy for obsessive-compulsive disorder guided by a computer or by a clinician compared with relaxation as a control. *Journal of Clinical Psychiatry, 63*, 138–145.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry, 23*, 56–62.

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: methods for reporting variability and evaluating clinical significance. *Behavior Therapy, 15*, 336–352.

Jones, S. (2004). Psychotherapy of bipolar disorder: a review. *Journal of Affective Disorders, 80*, 101–114.

Lam, D. H., Hayward, P., Watkins, E. R., Wright, K., & Sham, P. (2005). Relapse prevention in patients with bipolar disorder: cognitive therapy outcome after 2 years. *American Journal of Psychiatry, 162*, 324–329.

Lam, D. H., Jones, S. H., Hayward, P., & Bright, J. A. (1999). *Cognitive therapy for bipolar disorder: A therapist's guide to concepts, methods and practice*. Chichester: Wiley.

Lam, D. H., Watkins, E. R., Hayward, P., Bright, J., Wright, K., Kerr, N., et al. (2003). A randomised controlled study of cognitive therapy for relapse prevention for bipolar affective disorder. *Archives of General Psychiatry, 60*, 145–152.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174.

Leslie, J. C., & O'Reilly, M. F. (1999). *Behavior analysis: Foundations and applications to psychology*. Netherlands: Harwood Academic Publishers.

Mansell, W. (2006). The Hypomanic Attitudes and Positive Predictions Inventory (HAPPI): a pilot study to select cognitions that are elevated in individuals with bipolar disorder compared to non-clinical controls. *Behavioural and Cognitive Psychotherapy, 34*, 467–476.

Mansell, W. (2007). An integrative formulation-based cognitive treatment of bipolar disorders: application and illustration. *Journal of Clinical Psychology, 63*, 447–461.
Mansell, W. (2010). Bipolar disorder. In A. Grant, M. Townend, R. Mulhern, & N. Short (Eds.), *Cognitive behavioural therapy in mental health care* (2nd ed.). (pp. 142–164) London: Sage Publications.
Mansell, W., & Hodson, S. (2009). Imagery and memories of the social self in people with bipolar disorders: empirical evidence, phenomenology, theory and therapy. In L. Stopa (Ed.), *Imagery and the threatened self: Perspectives on mental imagery and the self in cognitive therapy* (pp. 232–264). London: Routledge.
Mansell, W., & Jones, S. H. (2006). The Brief-HAPPI: a questionnaire to assess cognitions that distinguish between individuals with a diagnosis of bipolar disorder and non-clinical controls. *Journal of Affective Disorders, 93*, 29–34.
Mansell, W., & Lam, D. (2006). "I won't do what you tell me!": elevated mood and the assessment of advice-taking in euthymic bipolar I disorder. *Behaviour Research and Therapy, 44*, 1787–1801.
Mansell, W., Morrison, A. P., Reid, G., Lowens, I., & Tai, S. (2007a). The interpretation of, and responses to, changes in internal states: an integrative cognitive model of mood swings and bipolar disorders. *Behavioural and Cognitive Psychotherapy, 35*, 515–539.
Mansell, W., Morrison, A. P., Reid, G., Lowens, I., & Tai, S. (2007b). An integrative cognitive model of mood swings and bipolar disorders: an illustrative case study. *Behavioural and Cognitive Psychotherapy, 35*, Supplementary materials.
Mansell, W., Paszek, G., Seal, K., Pedley, R., Jones, S. A., Thomas, N., et al. (2011). Extreme appraisals of internal states in bipolar I disorder: a multiple control group study. *Cognitive Therapy Research, 35*, 87–97.
Mansell, W., Powell, S., Pedley, R., Thomas, N., & Jones, S. A. (2010). The process of recovery from bipolar I disorder: a qualitative analysis of personal accounts in relation to an integrative cognitive model. *British Journal of Clinical Psychology, 49*, 193–215.
Marks, I. (1986). *Behavioural psychotherapy*. Bristol, UK: John Wright. Now published by I. Marks, Institute of Psychiatry, London.
Morrison, A. P. (2001). The interpretation of intrusions in psychosis: an integrative cognitive approach to hallucinations and delusions. *Behavioural and Cognitive Psychotherapy, 29*, 257–276.
Mundt, J. C., Clarke, G. N., Burroughs, D., Brenneman, D. O., & Griest, J. H. (2001). Effectiveness of antidepressant pharmacotherapy: the impact of medication compliance and patient education. *Depression and Anxiety, 13*, 1–10.
Mundt, J. C., Marks, I. M., Shear, M. K., & Griest, J. H. (2002). The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *British Journal of Psychiatry, 180*, 461–464.
National Institute for Health and Clinical Excellence. (2006). *The management of bipolar disorder in adults, children and adolescents, in primary and secondary care*. London: NICE.
Neff, K. D. (2003). The development and validation of a scale to measure self-compassion. *Self and Identity, 2*, 223–250.
Powers, W. T. (1973). *Behaviour: The control of perception*. Oxford, UK: Aldine.
Rice, M. J. (2008). Evidence-based practice in psychiatric care: defining levels of evidence. *Journal of the American Psychiatric Nurses Association, 14*, 181–187.
Scott, J., Paykel, E., Morriss, R., Bentall, R., Kinderman, P., Johnson, T., et al. (2006). Cognitive–behavioural therapy for severe and recurrent bipolar disorders. *British Journal of Psychiatry, 188*, 313–320.
Simon, N. M., Otto, M. W., Wisniewski, S. R., Fossey, M., Sagduyu, K., Frank, E., et al. (2004). Anxiety disorder comorbidity in bipolar disorder patients: data from the first 500 participants in the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). *American Journal of Psychiatry, 161*, 2222–2229.
Udachina, A., & Mansell, W. (2007). Cross-validation of the Mood Disorders Questionnaire, the Internal State Scale, and the Hypomanic Personality Scale. *Personality and Individual Differences, 42*, 1539–1549.
Watkins, E., Scott, J., Wingrove, J., Rimes, K., Bathurst, N., Steiner, H., et al. (2007). Rumination-focused cognitive behaviour therapy for residual depression: a case series. *Behaviour Research and Therapy, 45*, 2144–2154.
Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070.
Wells, A., & Sembi, S. (2004). Metacognitive therapy for PTSD: a preliminary investigation of a new brief treatment. *Journal of Behavior Therapy and Experimental Psychiatry, 35*, 307–318.