



Special issue paper

The effectiveness of cognitive analytic therapy for borderline personality disorder: Utilizing a withdrawal experimental design to improve sensitivity to abandonment

Stephen Kellett^{1*} , Joe Gausden² and Chris Gaskell²

¹University of Sheffield and Sheffield Health and Social Care NHS Foundation Trust, UK

²University of Sheffield, UK

Objectives. A primary methodological weakness of the single-case experimental design (SCED) outcome studies conducted of the treatment of personality disorder with cognitive analytic therapy (CAT) is that they have failed to employ a withdrawal phase or cross-over design and so are deemed quasi-experimental. This study sought to implement a withdrawal design, in order to improve the internal validity of the study to make it a true SCED, and also in order to enable the patient to explore abandonment dynamics.

Design. The study employed an A¹/B¹/A²/B² with extended follow-up SCED with a female patient treated with CAT meeting diagnostic criteria for borderline personality disorder (BPD). Following the 6-week baseline period 'A¹', treatment occurred in two phases (21 sessions 'B¹' and 24 sessions 'B²') sandwiching a 12-week treatment withdrawal phase ('A²') and a 24-week structured follow-up phase. Seven idiographic daily measures were collected that created a $N = 698$ day timeline. Nomothetic outcome measures were collected at baseline and at the end of each phase of the study, and the Session Impact Scale was completed after each treatment session.

Results. There was a significant increase in the task focus of treatment sessions. Ideographically, CAT was an effective treatment for improving the participant's self-to-self relationship, as their self-hate reduced and their sense-of-self increased. There was a broad pattern of deterioration during the second treatment phase (B²) and follow-up phase across the ideographic measures, and CAT was ineffective for BPD ideographic emotional or self-to-other measures. Reliable change occurred on the primary BPD nomothetic outcome measure from baseline to end of first treatment phase.

Conclusions. The study suggests that the CAT intervention was partially successful and that it is possible to integrate good research practice with clinical innovation. The methodological strengths and limitations of the design and the clinical implications of the results are discussed.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

*Correspondence should be addressed to Stephen Kellett, Clinical Psychology Unit, Department of Psychology, University of Sheffield, Sheffield S10 2TN, UK (email: S.Kellett@sheffield.ac.uk).

Practitioner points

- Cognitive analytic therapy was partially effective in the treatment of BPD in a case that had been unresponsive to other psychological interventions.
- Therapists need to complete lengthy and structured follow-up to capture any emerging relapse.
- Therapists need to discuss the patient's thoughts and feelings about termination regularly in a relationally informed manner.

The DSM-5 (APA, 2013) dimensional trait model of personality disorder (PD) consists of five broad domains of negative affectivity, detachment, psychoticism, antagonism, and disinhibition. As noted, 'these five broad domains are maladaptive variants of the five domains of the extensively validated and replicated personality model known as the 'Big Five,' or the Five Factor Model of personality' (APA, 2013, p. 773). Borderline personality disorder (BPD) can therefore be characterized using the five-factor model in terms of the personality features of persistent dysregulated rage, despondency, self-disturbance, emotional dysregulation, behavioural dysregulation, dissociation, and impulsivity (Widiger, 2015). BPD patients can be a challenge to treat psychotherapeutically as chronically dysfunctional interpersonal relationships (Widiger, 2015) can get repeated within and so negatively impact on the therapeutic dynamics, therefore creating pressure on boundary management in the patient–therapist relationship (Fonagy & Bateman, 2006; Skodol, *et al.*, 2002). Such relational difficulties mean that patients with PD can often face the double stigma of having a complex and enduring mental health problem and also being part of often difficult, unhelpful and potentially iatrogenic interactions with services (Black *et al.*, 2011). BPD patients are therefore high consumers of resources across in-patient, community team, and primary care settings, but where staff can experience BPD patients as 'hard-to-help' or 'difficult' due to their often provocative, confusing and alienating interpersonal behaviours (Kerr, 1999).

Clearly, BPD is both a contested and stigmatized psychiatric diagnosis, particularly when used with younger people (Koehne, Hamilton, Sands, & Humphreys, 2013), and so BPD generates more research and associated debate than all the other personality disorders (Boschen & Warner, 2009). The BPD diagnosis is frequently critiqued as a construction which pathologizes any deviation from societal norms (Shaw & Proctor, 2005), and this is particularly apparent with female patients (Wirth-Cauchon, 2001). BPD and bipolar disorder (types I and II) are also frequently confused due to the apparent mood-related symptomatic overlap, with such misdiagnosis then preventing any potentially effective intervention being delivered (Paris & Black, 2015). The cognitive analytic approach to BPD neither dismisses nor supports the diagnostic debate, but rather presents a structural model of BPD that focuses on the presence of multiple self-states (Pollock, Broadbent, Clarke, Dorrian, & Ryle, 2001). These self-states are established as a means of surviving childhood trauma and are maintained through ongoing dissociation, with three levels of increasing damage to reflective capacity possible (Ryle, 1997). The self-states approach enables the CAT therapist to name, map and work with separate states in an effort to enable integration, rather than being confused or over-whelmed by the multiplicity of apparent problems (and associated means of relating) the patient might present with both across and within sessions (Ryle & Kellett, 2018).

Cognitive analytic therapy (CAT) as a brief, integrative and relational psychotherapy (Ryle & Kellett, 2018) appears to be an acceptable, safe and effective intervention for BPD, as evidenced by the (admittedly small) number of cohort and randomized controlled trials completed thus far (Chanen *et al.*, 2008; Kellett, Bennett, Ryle & Thake, 2013; Ryle & Golyunkina, 2000; Wildgoose, Clarke & Waller, 2001). These studies use the mean group

response to analyse change, which unfortunately obscures the responsivity (or not) of individual patients (Heneghan, Goldacre, & Mahtani, 2017). Therefore, as a supplement to evidence generated by cohort and controlled studies, single-case experimental designs (SCED) are a potentially useful and cost-effective method of determining specific mechanisms of change in psychotherapies for BPD (Rizvi & Nock, 2008). SCEDs focus on individual patients aiming to predict and influence therapeutic change via the defining features of (1) repeated and intensive sampling of ideographic measures starting with a baseline, (2) manipulation of one or more independent variables whilst controlling for sources of bias, and (3) demonstration of stability within and across levels of imposed independent variables (Berghoff & Forsyth, 2015). The more controlled the SCED methodology, the more likely the identification of important facilitators of change and/or insights into the role played by clinically relevant variables (Barlow, Nock & Hersen, 2009).

Studies have investigated BPD outcomes via various SCED methods for behavioural therapy (Bloxham, Long, Alderman, & Hollin, 1993), cognitive therapy (Davidson & Tyrer, 1996), schema therapy (Nordahl & Nysæter, 2005), and dialectical behaviour therapy (Rizvi & Linehan, 2005). There have however been no previous SCED studies of CAT for BPD. There is however a CAT SCED evidence base with other personality disorders: paranoid personality disorder (PPD; Kellett & Hardy, 2014), histrionic personality disorder (HPD; Kellett, 2007), and dependent personality disorder (Kellett & Lees, 2019). It is worth noting that all of these studies used the basic bi-phasic A/B design and therefore would be categorized as only quasi-experimental (Shadish & Sullivan, 2011) – despite the lengthy additional follow-up phase in each of these studies. An A/B case series would be seen as experimental when participants are randomized to a staggered treatment start in a multiple baseline design (Christ, 2007). In an $N = 1$ study, in order to evaluate outcomes with greater confidence therefore requires some form of experimental manipulation, such as a withdrawal of treatment or a treatment comparison (i.e., cross-over) phase (Barlow, Nock & Hersen, 2009).

The present study therefore used a true SCED design single case that involved a treatment withdrawal phase, an A/B/A/B design (Hersen, 1990). An A/B/A/B design measures a baseline phase in ideographic measures (A^1 in the current study), a treatment phase (the B^1 in the current study), the withdrawal of treatment (A^2 in the current study), and the reintroduction of treatment (the B^2 in the current study). Effective interventions in such designs are recorded when the removal of treatment (A^2) means that outcomes return to baseline levels (A^1), and this is most readily implemented with purely behavioural interventions. It is acknowledged that because the current study was evaluating a psychotherapy, where the aim/hope is to facilitate learning in the patient, then the removal of treatment is not as ‘clean’ as during purely behavioural interventions. Therefore, any change that occurs during the first treatment phase (B^1) may be retained across the treatment removal phase, due to the patient internalizing the change methods in the first treatment phase. It is recognized that there is also a marked ethical dilemma when conducting withdrawal designs, as the design is beneficial from a scientific point of view, but potentially harmful to the patient, particularly when previously effective treatments are removed (Hersen, 1990). Treatment withdrawal has the potential for cuing deterioration; however, such a design is essential in the reliable attribution of progress to treatment effects (Hersen, 1990).

There have been previous calls for the employment of withdrawal designs in the CAT SCED literature (Kellett & Lees, 2019). Clinically, the withdrawal design was also indicated due to a fear of abandonment or rejection sensitivity being key feature of BPD

(Palihawadana, Broadbear & Rao, 2019). Rejection sensitivity is defined as ‘the disposition to anxiously expect, readily perceive and intensely react to rejection’ (Downey, Mougios, Ayduk, London, & Shoda, 2004, p. 668), and in the CAT model, this would be summarized as an *abandoning–abandoned* reciprocal role (Ryle & Kerr, 2003). Patients with BPD often struggle with such abandonment issues during and at the end of therapy, and CAT tries to anticipate this in the narrative reformulation and also the goodbye letters (which are distinctive features of the CAT approach; Ryle & Kellett, 2018) and in the analysis of abandoning–abandoned reciprocal role dynamics within the therapeutic relationship and in relationships with others, such as staff involved in the care of the patient (Ryle, Kellett, Hepple, & Calvert, 2014). Therefore, the opportunity of an A/B/A/B design was that the treatment withdrawal (i.e., A²) would allow the participant to practise tolerating a period without therapy, test out abandonment fears, express feelings about the ending, enable the analysis of associated enactments in the therapeutic relationship, and also give an opportunity to put into place the ‘exits’ learnt during the first phase of treatment. Palihawadana *et al.* (2019) have particularly called for studies investigating the role of abandonment fears during the treatment of BPD. The hypotheses for the current study were as follows: (1) there would be significant improvements in idiographic BPD measures during active treatment phases compared to baseline and withdrawal, (2) changes in the ideographic BPD measures would be sustained over the follow-up period, and (3) there would be a clinical and reliably significant change on the primary nomothetic BPD outcome measure.

Method

Design

The reporting of this study is based on the single-case reporting guidelines (SCRIBE; Tate *et al.*, 2016), and ethical approval for the study was granted (ref: 032506). The participant provided consent for the study to be conducted and reported, and this is consistent with guidance on the ethics of reporting single cases (Cooper, Turpin, Bucks, & Kent, 2005). The study uses an A/B/A/B design, with an extended follow-up (Hersen, 1990) and contained a range of ideographic and nomothetic measures. The study employed a structured 6-month follow-up period, as that is part of the treatment model of CAT with BPD (Ryle & Kellett, 2018). Seven idiographic measures were completed on a daily basis throughout all phases of the study. The baseline phase (A¹) spanned 6 weeks and three sessions. Treatment phase (B¹) lasted for 24 weeks and contained 19 treatment sessions. The treatment withdrawal phase (A²) spanned 16 weeks. The second treatment phase (B²) spanned 28 weeks and consisted of 24 sessions. The follow-up phase was 25 weeks. The study therefore constituted a time series of $N = 698$ days containing 5 distinct phases. Nomothetic outcome measures were completed at start of the baseline phase and then at the end of each phase of the study. The Session Impact Scale (Elliott & Wexler, 1994) was administered after each treatment session (i.e., 48 sessions). The SIS scores were categorized as baseline (sessions 1–3 during A¹), 4–24 (i.e., treatment sessions during B¹), and 25–48 (i.e., treatment sessions during B²).

The patient

The patient was 46-year-old single woman, who lived alone in a rented flat and was unemployed and claiming associated welfare benefits. The patient had been previously

assessed by a psychiatrist, and the diagnosis was BPD with comorbid recurrent depressive disorder. Before the start of the present study, the patient was interviewed with the Structured Clinical Interview for DSM-IV (SCID-I/SCID-II; First, Gibbon, Spitzer, Williams, & Benjamin, 1997). This semi-structured clinical interview yields a diagnosis consistent with DSM-IV/DSM-IV-TR (American Psychiatric Association, 2000) diagnostic criteria, and the patient met diagnostic criteria for BPD. The patient had been in contact with mental health services for 18-years prior to the present study. Throughout contact with services, the patient complained of persistent low mood, poor sleep, hopelessness, isolation, active suicidal thoughts, problems with impulse control, hearing a critical female voice, compulsive exercise, substance misuse, self-harm, and interpersonal intolerance. The patient was consistently described as very difficult to engage by services. The patient was referred by a community psychiatric nurse (CPN) whom was struggling to contain and help her due to her state variability, impulsive acts of self-harm, and frequent requests for help, only to be dissatisfied with the help then being offered.

Throughout the present study, the patient was prescribed two antidepressants (Mirtazapine and Venlafaxine) and also took a tranquilliser (Zopiclone) when her sleep was poor. She had a previous course of counselling in primary care, but reported feeling patronized by the counsellor and dropped out. The patient dropped out of CBT offered through the Improving Access to Psychological Therapies (IAPT) service. She was in contact with the community mental health team at the time of the study and had a 6-month psychiatric review and infrequent home visits from a community nurse. The patient had attended a psychoeducational group regarding BPD management which she found helpful.

The patient was raised in a nuclear family with two female siblings. The relationship with the father was described as poor as a child due to his violence and one of the grandparents was particularly verbally critical. She was chronically bullied at school, was sexually abused as a child, and was raped aged 16 years, resulting in her withdrawing from the world and relationships. There was a period of heavy amphetamine and MDMA abuse in her early twenties. Her first marriage lasted for 4 years, but was marred by violent rage attacks on her husband. Her next relationship was also marred by her violence to her then partner. The next significant relationship resulted in a daughter, but was terminated after a very violent assault on her by the partner. Physical assault was the norm in this relationship. The patient fled and temporarily lived in a women's refuge. The second marriage was more harmonious, but the patient stated that she never loved her husband and suddenly decided to leave him. The patient described a very limited life in which she avoided the social world and distrusted people in general. She would spend her time compulsively exercising (up to 3 hr per day) in the flat as a means of managing her mood and the voices that she heard. When she was particularly low in mood, she reported hearing two female voices, one of which was very derogatory and negative. The patient would have periodic substance misuse binges. She stated that she had little idea who she was as a person and could not articulate any personal values. She reported a sense of personality fragmentation and quickly altering between self-states. The patient reported a deep distrust of others and noted that she assumed that if she allowed someone to get close to her, then they would hurt and/or abandon her. The patient reported the presence of dissociative episodes and trance states. She gave numerous examples of exiting dissociative episodes and finding the damage that she had done to her property. The patient had not worked for many years due to presence of psychological symptoms. Previous employment experiences had tended to flounder because of aggressive behaviour

towards colleagues. She described her mood as typically bleak and flat. She stated that she hated herself and that she was disgusting.

Treatment

Treatment was delivered in the United Kingdom in a tertiary outpatient psychotherapy service provided by the National Health Service. The therapist was a male Consultant Clinical Psychologist and CAT psychotherapist and had clinical supervision provided by a UKCP CAT psychotherapist. Following a screening session, the patient was allocated to two blocks of 24 sessions of CAT, plus six-month follow-up. Treatment in CAT is normally 8, 16, or 24 sessions, but extended treatment contracts for BPD are possible according to the needs of the patient (Ryle & Kerr, 2003). The treatment plan was to do specific trauma-focused work in the second block of treatment after the stabilization of the first phase. The patient had a poor response to the start of the trauma treatment and deteriorated rapidly – therefore the second block of the CAT continued the work of the first. The rationale for having a treatment withdrawal phase to the study was explained to the patient as that this period would allow her to practise the exits learnt in the first phase of therapy and also practise the therapy coming to an end and therefore being able to process and explore abandonment issues. Sessions were weekly and lasted for 50 min. All sessions were attended and were conducted by the same therapist regardless of study phase. The first three sessions were focal to assessment tasks (e.g., taking a history etc.) and so did not contain any treatment elements and culminated in a narrative reformulation (i.e., read at session four). The narrative reformulation made links between the past and present, and stated the target problems and target problem procedures (Ryle & Kellett, 2018). This was considered the end of the baseline and the start of active treatment, as is consistent with previous CAT SCED research (Kellett, Simmonds-Buckley & Totterdell, 2016).

In terms of treatment fidelity, the hallmark components of CAT therapy are a narrative reformulation, a sequential diagrammatic reformulation, and goodbye letters exchanged at the termination of therapy by patient and therapist (Ryle & Kerr, 2003); in the current case, all these distinctive features of the CAT model were present. Each component was reviewed and appraised at clinical supervision, and sections of audio-recording from sessions were taken to supervision. One entire 50-min audio-recorded treatment sessions were sampled from each active phase of treatment (sessions 16 in phase 1 and session 34 of phase 2). These were rated by the therapist with the Competence in Cognitive Analytic Therapy measure (Bennett & Parry, 2004). The CCAT contains 77 elements of therapist competence across 10 domains of therapeutic practice. The CCAT score was 30/40 for session 16 and 32/40 for session 34, with 20 being the cut-off for competent CAT (Bennett & Parry, 2004).

Treatment was theoretically grounded in the sequential diagrammatic formulation which was underpinned by CAT's multiple self-states model (MSSM; Pollock *et al.*, 2001) describing the following key states: 'blind rage' (with the reciprocal roles of abusing/attacking to abused/shamed/weak), 'lost-world' (with the reciprocal role of abandoning–abandoned), 'protector' (with the reciprocal role of protecting safe), and 'old-bagging' (with the reciprocal role of criticizing/humiliating to humiliated/put down) and a numb cut-off dissociated state. The states were elicited using the states description procedure (SDP) approach (Ryle, 2007), and a self-states sequential diagrammatic reformulation was co-produced with the patient (Ryle, Beard, & Marlowe, 1995). The 'exits' that the patient developed were (1) assertiveness, (2) walking away from rage, (3) stopping rescuing, (4) allowing the daughter to grow up, (5) limiting exercise to one-hour every two days, (6)

limiting alcohol intake, (7) being less self-critical and more cognitively flexible, (8) trusting self, and (9) exposure to social contact. CAT is also a therapy that actively works with enactments in the therapeutic relationship (Ryle & Kellett, 2018). The most common enactment was analysis of the 'old-bagging' state during treatment sessions and the 'lost-world' state during preparation for the treatment withdrawal phase and the end of the second active treatment phase.

Idiographic outcome measures

The seven idiographic measures were separated into three categories: emotions, self-to-self, and other-to-self. Each measure was scored on a 10-point Likert scale from 0 (not at all) to 10 (completely). The 'emotions' category contained three measures: *feeling abandoned*, *feeling anxious*, and *feeling lonely*. The 'self-to-self' category contained two measures: *bating myself* and *a poor sense-of-self*. The 'self-to-other' category contained two measures: *distrusting of others* and *over-sensitive*. Positive change was indicated by a decrease in scores in all idiographic measures, except 'sense-of-self' measure, where positive change was indicated by a score increase.

Nomothetic outcome measures

The Borderline Symptom List (BSL-23; Bohus *et al.*, 2009) assesses intensity of BPD symptoms over the previous week and contains 23 items (rated 0–4). BSL-23 caseness cut-off is a score of 2. The BSL-23 has good internal ($\alpha = .96$) and test-retest ($r = .82$, $p < .001$) reliability. The BSL-23 was the primary nomothetic outcome measure for the study. Beck Depression Inventory-II (BDII; Beck, Steer, & Brown, 1995); this measure contains 21 items that measure the intensity of depressive symptom (Beck, Steer, Ball & Ranieri, 1996), with a clinical cut-off score of 17. BDI-II scores are coded as 0–13 = minimal depression, 14–19 = mild depression, 20–28 = moderate depression, and 29–63 = severe depression. Brief Symptom Inventory (BSI; Derogatis, 1993); the BSI (53 items) is a valid and reliable measure of psychological distress (Derogatis, 1993) consisting of three subscales; the global severity index (GSI) is the most commonly reported. A raw score $>.78$ on the BSI-GSI relates to the patient reaching 'caseness' on the measure. Inventory of Interpersonal Problems-32 (IIP-32; Barkham, Hardy, & Startup, 1996); this tool measures interpersonal problems and is a shortened version of the IIP-126 (Barkham *et al.*, 1996). The IIP-32 has four scales relating to interpersonal functioning and four scales identifying problematic and dysfunctional interpersonal strategies (Hughes & Barkham, 2005).

Session impact measure

The Session Impact Scale reliably measures the experienced personal impacts of psychotherapy sessions in terms of the key features of the session (Elliott & Wexler, 1994). The 16 SIS items form three factors: the task focus of the session, the quality of the therapeutic relationship, and whether anything hindering occurred. Item 17 is used to provide qualitative data on any 'other important impacts' that occurred during the session.

Table 1. Summary of the function of the tests used in the analyses

Ideographic or nomothetic test	Function of the test
Ideographic; analysis of covariance (ANCOVA)	To examine the phase differences in the mean values of the ideographic measures that are related to the effect of the CAT intervention, whilst taking into account the influence of the uncontrolled independent variables
Ideographic; bonferroni correction	A multiple-comparison correction was used when several statistical tests are being performed simultaneously, that is employed in order to avoid producing spurious positive results
Ideographic; percentage exceeding the median	The analysis of the effectiveness of the CAT intervention phase data that is contingent on the overlap with the median data point within the baseline phase and is based on the assumption that if the CAT intervention (B phases) was effective, ideographic outcome data will predominately exist on the therapeutic side of the median; if the CAT intervention was ineffective, data points during the treatment B phases would vacillate above and below the baseline median
Ideographic; non-overlap of pairs	Calculation of non-overlap between the baseline and successive CAT intervention B phases; this test identifies the highest data point in baseline to determine the percentage of data points during CAT intervention B phases that exceeded this level and therefore assesses effectiveness
Ideographic; Tau-U	Tau-U statistics assessed change between phases whilst accounting for potential baseline trend. Tau-U analyses included (1) analysis of baseline trend, (2) difference between phases (τ^A vs. τ^B) which essentially emulates a Mann–Whitney test, and (3) $\tau^{(A-B)-Arend}$ which compares phases whilst adjusting for baseline trend. It is recommended that if baseline trend is not significant then $\tau^{(A-B)-Arend}$ is not indicated and that τ^A vs. τ^B should be used (Bossart, Laird, & Armstrong, 2018)
Nomothetic; reliable change	To differentiate change in a nomothetic outcome that is statistically significant from change that may have occurred due to measurement error
Nomothetic; reliable and clinically significant change	To identify when the change is both reliable (see above) and also when the scores have shifted from in the clinical range to the community range

Table 2. Mean (SD) of ideographic measures over the study phases

Ideographic daily measure (1–10) scale	Baseline mean (SD)	CAT first treatment phase mean (SD)	Treatment withdrawal (SD)	CAT second treatment phase mean (SD)	Follow-up (SD)	F value
Feeling abandoned	8.50 (0.75)	8.09 (1.02)	8.17 (0.99)	7.60 (1.39)	9.04 (6.06)	4.14**
Feeling anxious	8.73 (1.07)	8.16 (1.31)	8.09 (1.06)	7.92 (1.65)	8.86 (1.12)	3.79**
Feeling lonely	9.15 (0.85)	8.44 (1.24)	7.44 (0.88)	7.97 (1.53)	8.58 (1.73)	3.91**
Self-hate	9.41 (0.81)	8.70 (1.57)	6.18 (1.49)	6.34 (2.21)	8.32 (1.14)	11.24**
Sense-of-self	0.83 (0.97)	2.12 (1.37)	2.87 (0.94)	3.06 (2.01)	1.61 (0.94)	6.92**
Distrusting of others	8.98 (0.61)	9.24 (0.51)	8.82 (0.51)	8.92 (0.91)	9.57 (0.50)	6.33**
Feeling over-sensitive	9.56 (0.64)	9.05 (0.88)	8.56 (0.67)	8.84 (0.92)	9.49 (0.77)	4.90**

Note. ** $p < .001$.

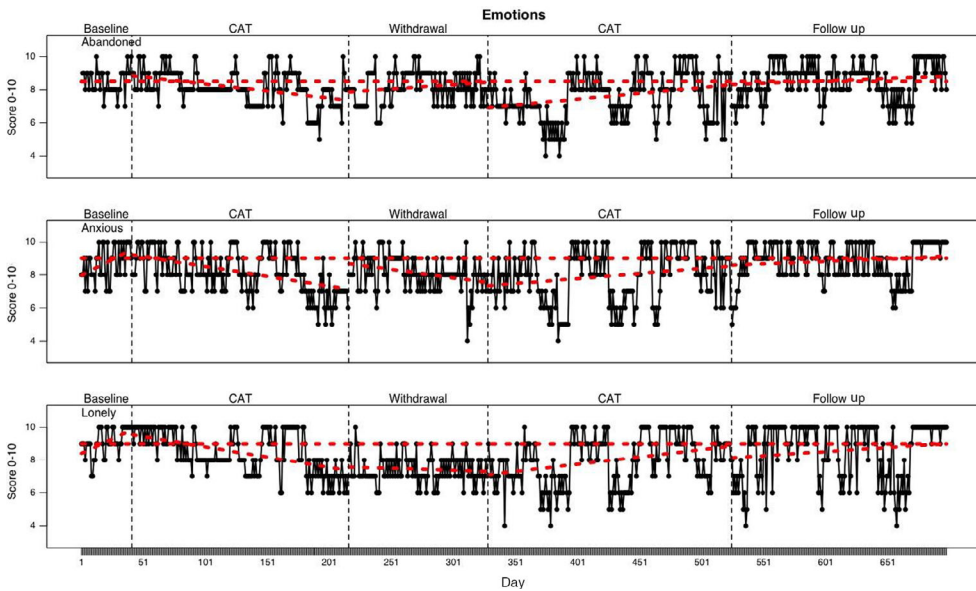


Figure 1. Time-series plots for ideographic emotion measures including baseline median (across phase horizontal dotted red line) and phase trend lines (within phase dotted red lines).

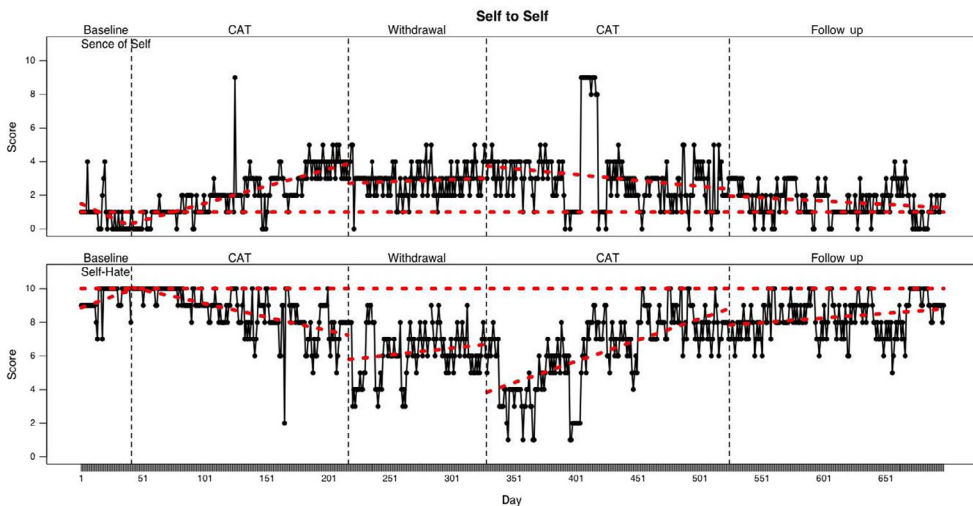


Figure 2. Time-series plots for self-to-self ideographics including baseline median (across phase horizontal dotted red line) and phase trend lines (within phase dotted red lines).

Analysis strategy

Table 1 provides a description of the tests used for the ideographic and nomothetic data and definitions of the purpose of the tests. To adjust for any autocorrelation in the time-series ideographic measures, then partial autocorrelation function (PACF) plots were calculated in order to show which lag in each ideographic measure was appropriate for subsequent use in the ANCOVA analysis. Autocorrelation is the non-independence of

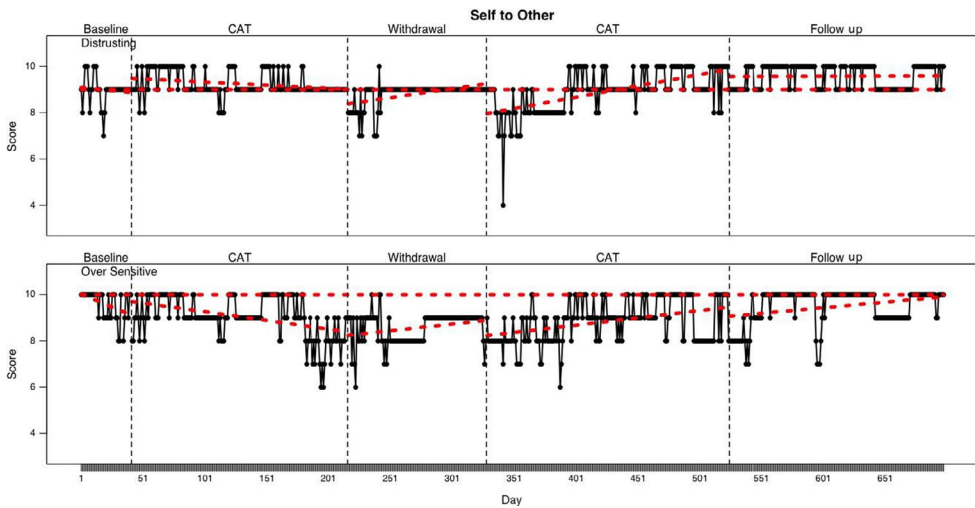


Figure 3. Time-series plots for self-to-other ideographics including baseline median (across phase horizontal dotted red line) and phase trend lines (within phase dotted red lines).

sequential observations in a time series and also known as serial dependence (Borckardt, Nash, Murphy, Moore, Shaw, & O'Neil, 2008). Autocorrelation in time-series data biases the standard errors and so potentially positively distorts subsequent results (Drucker, 2003), and so removing the autocorrelation ensured the reliability of the standard errors used in the ANCOVA. The PACF plots are displayed in the Figure S1. First-order lags were appropriate to be applied for all ideographic measures, except the feeling abandoned measure, where a second-order lag was most appropriate. ANCOVA then tested for any differences between the phases of the study, and the lags were used as covariates (Kellett & Totterdell, 2013). Despite the ideographic measures violating assumptions of homogeneity of variance, the analyses were still performed due to the robustness of ANCOVA (Schmider, Ziegler, Danay, Beyer, & Bühner, 2010). The ANCOVA had a single factor (study phase) and five levels ($A^1/B^1/A^2/B^2$ and FU), and a bonferroni correction was applied. Partial η^2 was used to calculate the effect sizes from the ANCOVAs and was interpreted as follows: 0.01 = small effect, 0.06 = medium effect, and 0.14 = a large effect. Analysis of non-overlap statistics (percentage exceeding the median; PEM and non-overlap of all pairs; NAP) was also performed in order to assess the degree of change between phases on ideographic measures. Time-series graphs (with trend lines fitted and baseline medians) were created for each ideographic measure and weekly SIS score to display change by phase of study. Changes in session impact were assessed via ANOVA of SIS subscale scores and also non-overlap statistics. Qualitative data from item 17 of each SIS score for 44 of 48 sessions that this item was completed were coded as either a positive, mixed, or negative impact (see Appendix S1 for the qualitative SIS comments), and these ratings were second-rated to assess inter-rater reliability; the percentage of agreement between the raters was 86.36%. Nomothetic outcomes were analysed using the reliable change index for each measure (RCI; Jacobson & Truax, 1991) and also reliable and clinically significant change (RCSC, Jacobson & Truax, 1991).

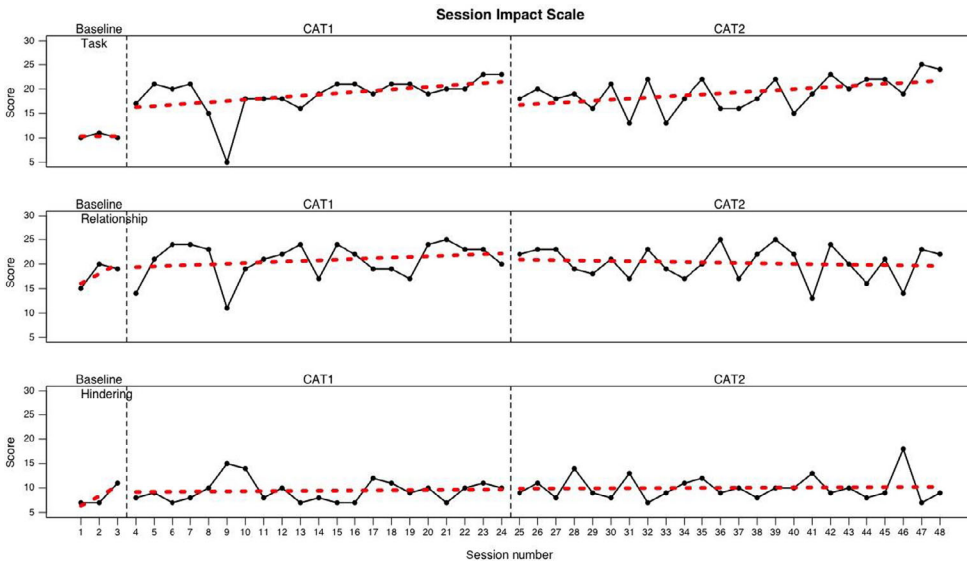


Figure 4. Plots for SIS subscales including phase trend lines (within phase dotted red lines) across baseline and first and second CAT treatment phases.

Results

Results are presented in three sections: idiographic outcomes (according to emotions, self-to-self, and self-to-other), nomothetic outcomes, and then session impact scores. The general patterns across the ideographic outcomes were that the first treatment phase created reductions in BPD ideographic outcomes, but there were patterns of deterioration evident over the follow-up period. Table 2 provides a summary of means (*SD*) on the ideographic measures reported by study phase, whilst Table 3 contains the non-overlap and TAU-U results comparing the phases. Where Tau baseline trend for an ideographic measure was significant, Tau-U statistics are reported, and where baselines trends were stable, Tau^{A-B} is reported. Figures 1–3 present the time-series graphs for each ideographic measure.

Idiographic BPD outcomes; emotions

Visually, all the emotion measures showed an improvement during the first treatment phase and withdrawal, there was a variable response to the second treatment phase, and all showed deterioration over the follow-up (see Figure 1). There was a significant effect of the phase for feeling abandoned with a small effect size; $F(4, 690) = 4.14, p < .05, \eta^2 = .02$. The non-overlap results indicated that CAT was ineffective (PEM = 61.41%, NAP = 59.47) when baseline was compared with all subsequent phases pooled together; however, this effect was statistically significant when analysed using Tau^{a-b} ($\tau = -.184, p < .05$). Tau^{a-b} was not significant when comparing baseline and follow-up alone. There was a significant effect of the phase for feeling anxious, with a small effect size; $F(4, 691) = 3.79, p < .05, \eta^2 = .02$. CAT was ineffective on anxiety when baseline was compared with all other phases combined (PEM = 58.50%; NAP = 58.86). $\text{Tau}^{(a-b)\text{-trend}}$ was significant between baseline and the first treatment phase ($\tau = -.248, p < .05$), and

Table 3. Non-overlap results comparing the ideographic measures between phases

Daily measure	Baseline vs. first CAT treatment phase (CAT 1)				First CAT treatment phase (CAT 1) vs. follow-up		First CAT treatment phase (CAT 1) vs. withdrawal		Baseline + withdrawal vs. CAT 1 + 2				
	Baseline trend		τ^A vs. B		τ -U	NAP	PEM	NAP	PEM	NAP	PEM		
Abandoned	-.002		-.224*		-.201*	61.48	67.82	69.33	54.91	59.29	58.89	59.09	34.77
Anxious	.278*		-.244*		-.248*	62.21	60.34	65.75	60.34	65.19	65.77	53.56	36.39
Lonely	.377*		-.304*		-.307*	65.18	55.17	62.51	62.64	89.99 ^a	88.29 ^a	42.10	31.00
Self-hate	.288*		-.339*		-.334*	66.97	68.39	77.82 ^a	76.44 ^a	96.75 ^b	100 ^b	43.26	27.22
Sense-of-self	-.326*		.574*		.548*	78.69 ^a	85.06 ^a	74.30 ^a	81.03 ^a	92.14 ^b	99.1 ^b	52.12	24.53
Distrusting	-.062		.211*		.195*	39.47	4.45	70.94 ^a	56.90	56.24	15.32	40.95	16.98
Over-sensitive	-.0340*		-.340*		-.281*	67.44	66.66	69.79	63.22	84.74 ^a	97.3 ^b	45.99	31.35
Baseline vs. second CAT treatment phase (CAT 2)													
Baseline vs. follow-up													
NAP	PEM	τ^A vs. B		τ -U	NAP	PEM	τ^A vs. B		τ -U	NAP	PEM	Baseline vs. CAT 1, withdrawal, CAT 2 and follow-up	
Abandoned	69.22		71.79 ^a		.068	.061	46.50	45.09	-.184*	-.179*	59.47	61.41 (38.41)	
Anxious	63.35		69.23		.073	.037	46.36	39.88	-.177	-.180*	58.86	58.5	
Lonely	71.81		57.44		-.061	-.090	53.06	37.57	-.364*	-.363*	68.18	56.81	
Self-hate	91.86		94.87 ^b		-.540*	-.514*	77.02 ^a	82.66 ^a	-.641*	-.631*	82.11 ^a	85.45 ^a	
Sense-of-self	87.15		96.41 ^b		.483*	.467*	74.14 ^a	88.44 ^a	.646*	.637*	82.30 ^a	91.73 ^b	
Distrusting	50.68		28.72		.486*	.442*	25.71	0	-.641*	-.631*	82.11 ^a	85.52 ^a	
Over-sensitive	72.40		70.26 ^a		-.034	.006	51.73	36.00	-.345*	-.328*	78.69 ^a	65.24	

Note. ^aIndicates highly effective treatment. ^bIndicates moderately effective treatment. *Significant at $p = .005$.

between baseline and all other phases was ($\tau = -.180, p = <.05$); however, $\text{Tau}^{(a-b)\text{-trend}}$ was not significant when comparing baseline and follow-up. There was a significant effect of phase on loneliness, with a small effect size; $F(4, 691) = 3.91, p < .05, \eta^2 = .02$. Non-overlap results indicated that CAT was ineffective for loneliness (PEM = 56.81%; NAP = 68.15%). $\text{Tau}^{(a-b)\text{-trend}}$ statistics were significant when comparing the loneliness baseline and the first CAT treatment phase ($\tau = -.307, p = <.05$), and between baseline and all other phases ($\tau = -.363, p = <.05$). However, the $\text{Tau}^{(a-b)\text{-trend}}$ was not significant when comparing baseline loneliness to follow-up loneliness.

Idiographic BPD outcomes; self-to-self

Visually, both self-to-self measures responded to the first treatment phase and also the withdrawal, but deteriorated over the follow-up (see Figure 2). There was a significant effect of phase on self-hate with a medium effect size; $F(4, 689) = 11.24, p < .001, \eta^2 = .06$. Non-overlap results indexed that CAT was moderately effective for self-hate (PEM = 85.45%; NAP = 82.11%). $\text{Tau}^{(a-b)\text{-trend}}$ statistics were significant when comparing the self-hate baseline and the first treatment phase of CAT ($\tau = -.344, p = <.05$), baseline and all other phases ($\tau = -.514, p = <.05$), and baseline with follow-up ($\tau = -.631, p = <.05$). There was a significant effect of phase on sense-of-self with a medium effect size; $F(4, 691) = 6.92, p < .001, \eta^2 = .04$. PEM results CAT to be a highly effective treatment for sense-of-self (PEM = 91.73%; NAP = 82.30%). $\text{Tau}^{(a-b)\text{-trend}}$ statistics were significant when comparing sense-of-self baseline and B1 ($\tau = -.548, p = <.05$), baseline and all other phases ($\tau = -.637, p = <.05$), and baseline with follow-up ($\tau = -.467, p = <.05$).

Idiographic BPD outcomes; self-to-other

Visually, the interpersonal over-sensitivity measure was stable during the first treatment phase and withdrawal, improved during the second phase, and deteriorated over the follow-up (see Figure 3). The distrust of others ideographic measure appeared fairly stable across all the phases (see Figure 3). There was a significant effect on interpersonal sensitivity, with a small effect size; $F(4, 687) = 4.90, p < .001, \eta^2 = .03$. Non-overlap results indicated that CAT was ineffective for over-sensitivity (PEM = 65.08%; NAP = 78.69%). $\text{Tau}^{(a-b)\text{-trend}}$ statistics were significant when comparing over-sensitive baseline and B1 ($\tau = -.281, p = <.05$), and between baseline and all other phases was ($\tau = -.328, p = <.05$); however, $\text{Tau}^{(a-b)\text{-trend}}$ was not significant when comparing baseline and follow-up. There was a significant effect on distrusting others, with a small effect size; $F(4, 691) = 6.33, p < .001, \eta^2 = .04$. Non-overlap results indicated an ineffective treatment for the ability to trust others (PEM = 85.51%; NAP = 81.11%). $\text{Tau}^{(a-b)}$ statistics were significant when comparing distrusting baseline and B1 ($\tau = -.211, p = <.05$), baseline and all other phases ($\tau = -.486, p = <.05$), and baseline and follow-up ($\tau = -.641, p = <.05$).

Nomothetic outcomes

Table 4 summarizes nomothetic outcomes. The BSL-32 showed reliable (but not clinical) change only from baseline to the end of the first treatment phase ($\text{RCI} = 2.11, p < .05$). On the BDI-II measure, reliable (but not clinical) change occurred only from the baseline phase to the end of the first treatment phase ($\text{RCI} = 3.55, p < .05$). The BSI-GSI measure

Table 4. Normothetic measure outcome (with normative comparisons)

Measure	Caseness cut-off	Baseline	CAT 1	Treatment withdrawal	CAT 2	Follow-up	Clinical mean (SD)	Community mean (SD)	RCI-1	RCI-2	RCI-3	RCI-4	Reliable and Clinically significant change from baseline to end of follow-up?
BSL-32	2.00	3.56	2.95	2.78	2.56	2.03	2 (0.76)	0.4 (0.22)	2.11*	0.59	0.76	1.84	No
BDI-II	17.00	51	31	28	25	27	20.44 (13.28)	10.04 (8.23)	3.55*	0.53	0.53	-0.35	No
BSI-GSI	63.00	3.35	2.65	2.24	1.98	2.03	1.32	0.30	2.17*	1.27	0.81	-0.16	Yes
IIP-32	1.39	(78.00)	(69.00)	(63.00)	(61.00)	(61.00)	(0.72)	(0.31)	-1.14	-0.61	1.92	-0.17	No
		1.71	2.12	2.34	1.65	1.71	1.51	0.98					

Note. RCI-1 is from the baseline to end of the first CAT treatment phase; RCI-2 is from the end of the first CAT treatment phase to the end of the treatment withdrawal phase; RCI-3 is from the end of the treatment withdrawal phase to the end of the second CAT treatment phase; RCI-4 is from the end of second CAT treatment phase to the end of follow-up. Scores in bold indicate where a score is in the clinical caseness range at that time.

*Indicates reliable change. Clinical and community norms were taken from the following papers: Beck Steer Ball and Ranieri (1996) for BDI-II, Derogatis (1993) for BSI-GSI, Barkham et al. (1996) for IIP-32, and Bohus et al. (2009) for BSL-32.

Table 5. Session impact results

Means (SD)				PEM (NAP)			
	Baseline M (SD)	CAT 1 M (SD)	CAT 2 M (SD)	F-value	Baseline sessions vs. CAT 1	Baseline sessions vs. CAT 2	Baseline vs. all CAT treatment sessions
SIS subscale							
Task	10.33 (0.58)	18.86 (3.79)	19.21 (3.27)	9.06*	95.24** (95.24**)	100 (100)	97.78 (97.78)
Relationship	18.00 (2.65)	20.76 (3.66)	20.25 (3.31)	0.86	66.67 (76.191)	62.50 (72.22)	64.44 (74.07)
Hindering	8.33 (2.31)	9.43 (2.27)	10.04 (2.51)	0.86	76.19 (65.08)	91.67 (72.22)	84.44 (68.89)
Item 17: Other important impacts							
Negative impact N (%)	1 (33%)	4 (19.05%)	6 (26.09%)				
Mixed impact N (%)	2 (66%)	3 (14.29%)	5 (21.74%)				
Positive impact N (%)	0	12 (57.14%)	9 (39.13%)				

Note. *Significant at $p = .005$; **Indicates highly effective treatment.

showed reliable ($RCI = 2.17, p < .05$) and clinically significant change from the baseline phase to the end of the first treatment phase. The BSI-GSI was in the non-caeness range at the end of the second phase of CAT treatment and at follow-up. On the IIP-32 measure, there was no reliable change across any phase of the study and also no clinically significant change.

Session impact outcomes

Table 5 reports the session impact scores on the SIS subscales and the qualitative item (i.e., item 17). Plots of session impact sub-domains can be seen in Figure 4. The task focus in the sessions significantly increased over the phases of the study, $F(2, 45) = 9.06, p = <.001$. There were high levels of non-overlap and hence impact when comparing baseline sessions with sessions during the first CAT treatment phase (PEM = 95.24%; NAP = 95.24%), baseline and the second CAT treatment phase (PEM = 100%; NAP = 100%), and baseline and combined CAT treatment phases (PEM = 97.78%; NAP = 97.78%). There was no significant change on the relationship scale, $F(2, 45) = 0.86, p = .43$ nor the hindering scale, $F(2, 45) = 0.86, p = .43$ over time. During the baseline phase sessions, the most common impact was mixed (2/3); during the first treatment phase; the most common session impact was positive (12/19); and during the second treatment phase, the most common session impact again was positive (9/24).

Discussion

The purpose of this study was to evaluate the effectiveness of CAT for a patient with BPD in a previously unused A/B/A/B design, and in particular to explore whether a treatment withdrawal phase could be matched to helping the patient better prepare for the actual end of the therapy, as fear of abandonment is a core feature of BPD (Palihawadana *et al.*, 2019). The length of the time series achieved is unique in the BPD SCED outcome evidence base regardless of treatment modality. The fidelity and competency evidence provided reassurance that CAT was delivered, the session impact scores were consistently high, and qualitative session impact themes tended to be positive. Four of the seven ideographic outcome measures had unstable baselines (i.e., feeling abandoned, self-hate, poor sense-of-self, and being over-sensitive), but it is worth noting that instability tended to reflect that deterioration in these measures was occurring during the baseline. In a disorder like BPD where variability is actually part of the clinical picture (Ryle & Kerr, 2003), then it is perhaps naive to expect stability during SCED baselines. The principle clinical concern in the current study was the observed deterioration from second treatment phase to follow-up, considering a main challenge with BPD patients is reducing over-dependence on mental health services (Nehls, 2000). Whilst consistent efforts were made to prepare the participant for the termination of therapy, the deterioration evident in the time-series data raises doubts about whether CAT (in this case) adequately prepared the participant. The participant made particularly good use of the treatment withdrawal phase and was not affected by any negative life events during this period. It was the impression of the therapist that the participant appeared to understand and appreciate the efforts being made to help her acclimatize to abandonment issues. The follow-up sessions were particularly marked by the participant struggling with their mood and some evidence of returning to habitual roles and patterns due to the series of negative life events.

Overall, the results suggest a partially effective intervention and previous clinical trial and cohort studies have found CAT to be effective in alleviating BPD symptoms (Chanen *et al.*, 2008; Kellett *et al.*, 2013). The partially effective outcome and also the good attendance of the participant need to be seen in the context of the previous poor therapy outcomes and rejection of previous psychological interventions. Where change occurred in the ideographic BPD measures, then it appears that the self-to-self measures were more responsive to treatment than the emotions or self-to-other measures. CAT appeared moderately effective in terms of reducing self-hate. This is an important reduction considering that suicidality and punishing self-injury among BPD patients often stem from feelings of self-hatred (Brown, Comtois & Linehan, 2002). The non-overlap results also suggested that CAT was a highly effective treatment for improving the 'sense-of-self' ideographic measure, and this is in line with previous evidence that CAT assists BPD patients in personality integration efforts (Kellett *et al.*, 2013). The evidence from this study is that that such personality integration work is both a lengthy and time-consuming effort, and perhaps BPD patients make a start with this goal during therapy and then spend a great length of time subsequent to the therapy continuing and completing this work.

The nomothetic outcomes of the study would mirror the ideographic outcomes in suggesting a partially effective intervention. In 3/4 nomothetic measures (BSL-32, BDI-II, and IIP-32), reliable change was seen from baseline to first treatment phase, but no further reliable change occurred. This evidence would be consistent with the 'dose-effect' relationship in psychotherapy, that the course of improvement over sessional time follows a negatively accelerating pattern of change, irrespective of actual eventual duration of the intervention (Robinson, Delgadillo, & Kellett, 2019; Rubel, Lutz, & Schulte, 2015). Just one nomothetic measure (BSI-GSI) recorded a realisable and clinically significant change from baseline assessment to the end of follow-up. It is worth noting that the BSL-32 was just above the cut-off score by the end of the follow-up. In terms of outcome, the 'task' subscale on the SIS significantly increased from baseline to treatment phases, indicating that treatment sessions were more task-focused. This increase in task focus is a feature of the phase change of CAT from reformulation to more active recognition/revision and has been previously illustrated (Spence, Kellett, Totterdell & Parry, 2019). Despite no significant change being recorded in the 'relationship' SIS subscale, scores were consistently high, indicating that throughout the therapy the patient felt positively towards the therapist and the course of therapy.

There are multiple possible explanations for the patient's mixed outcome. BPD patients commonly relapse and deteriorate after intervention (Jerschke, Meixner, Richter, & Bohus, 1998; Mohr, 1995) and individual differences in condition severity, chronicity, comorbidity with other conditions, and history of contact with mental health services all moderate how patients respond to the termination of interventions (Davidson & Scott, 2009). The participant in the current study had been in the mental health service for 18 years prior to the study commencing, had been non-responsive to previous psychological interventions, and was extremely socially isolated by choice. It has been previously argued that a reduction in distress during the early stages of CAT is crucial in enabling hope and before more complex state integration work is possible – but this does not seem to have been the case here (Kellett *et al.*, 2013). Given research on the effectiveness of CAT on BPD patients showing sub-syndromal or first-presentation symptoms (Chanen *et al.*, 2008), the patient's poor outcomes may be a result of late intervention (with a history of previously failed interventions), so illustrating the importance of addressing BPD early in its presentation.

The patient's attendance suggests the acceptability of CAT for BPD and also supports previous research showing CAT's consistently low drop-out rate (Calvert & Kellett, 2014). A methodological strength of the SCED was the introduction of a treatment removal phase and the length of the follow-up, with the daily idiographic measures indexing change on an intricate micro level. The collection of such an extensive time series suggests that the close alignment of the patient and the therapist on the design of the measures made them representative of the central goals of the work, and therefore, the measures were non-burdensome (Kellett & Beail, 1997). An A/B/A/B design possesses advantages over usual bi-phasic A/B designs (Kazdin, 1978), as repeated change in both treatment phases compared to non-treatment phases better tests that treatment has affected change, reducing the possibility of attributing change to other factors (passage of time, regression to the mean etc.; Rizvi & Nock, 2008). However, A/B/A/B designs are typically implemented with purely behavioural interventions and they are therefore not immune to 'bleed-over' in learning across the phases during psychotherapy (McMillan & Morley, 2010). Indeed, that was the aim of the intervention, to help the participant more effectively tolerate abandonment feelings through carryover of learning. Ethical dilemmas are also apparent when using treatment withdrawal phases for severe cases of BPD when the withdrawal of the treatment might dangerously increase self-destructive behaviours (Hersen & Barlow, 1976).

In terms of other limitations, the generalizability of the study to other BPD patients is open to question. The data were collected via self-report, which limits confidence in the reliability of the results, raising issues of social desirability bias (Arnold & Feldman, 1981; Nicklas, Dunbar & Wild, 2010). Therefore, some other informant data would have been useful (Kellett & Totterdell, 2013), but the participant's social isolation and fear of contact with others made this impossible. Additional nomothetic outcome measures may have been appropriate to use, for example, the Personality Structure Questionnaire (PSQ; Pollock, Broadbent, Clarke, Dorrian & Ryle, 2001) to assess change in state-shifting and the Scale for Suicide Ideation (e.g., the SSI; Beck, Kovacs & Weissman, 1979) to more closely capture change in risk issues. Wider sampling of competency across the treatment phases would have been useful also. The study methodology could have been improved through the addition of idiographic control and also generalization measures (Krasny-Pacini & Evans, 2018). But, a potential criticism is also the wide range of outcomes assessed and the associated risk of false-positive findings. This criticism is defended by the use of bonferroni corrections in the ANCOVA analysis of the ideographic outcomes. Finally, the benefits of the withdrawal design were research related, and so the repeat of such a design may not be as well tolerated by other BPD patients. It is worth noting that the treatment withdrawal phase was agreed at the outset, and therefore, treatment withdrawal was not 'sprung' on the patient, but rather diligently planned for. Future SCEDs with BPD patients may seek to utilize cross-over designs, as they appear to have less ethical dilemmas, due to random allocation to two different, but active, treatments (Jones & Kenward, 2014). For example, rather than withdrawing CAT, a BPD patient could be randomly allocated to CAT followed by CBT (or vice versa).

To conclude, this innovative and methodologically unique study has indexed a mixed outcome of CAT for a patient with BPD whom was treated in routine practice. The study has used a rigorous SCED methodology which heightens the reliability and internal validity of the study, and confidence in the conclusions drawn that the intervention was partially effective. The study's single-case design is the main challenge to the generalizability of the findings. As abandonment fears are central to BPD, the opportunity to use withdrawal designs is unique, but this needs to be carefully, clinically, and ethically

considered prior to implementation. This research nevertheless supports CATs continued use with BPD and also highlights the flexibility and rigour of the SCED method.

Conflicts of interest

All authors declare no conflict of interest.

Author contributions

Stephen Kellett: Supervision, Writing – original draft, Writing – review & editing. Joe Gausden: Data curation, Formal analysis, Writing – original draft. Christopher Gaskell: Formal analysis, Writing – review & editing.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author (SK). The data are not publicly available due to restrictions (e.g., their containing information could compromise the privacy of research participant).

References

- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders-IV-TR*. Washington, DC: American Psychiatric Association.
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.
- Arnold, H. J., & Feldman, D. C. (1981). Social desirability response bias in self-report choice situations. *Academy of Management Journal*, 24, 377–385.
- Barkham, M., Hardy, G. E., & Startup, M. (1996). The IIP-32: A short version of the Inventory of Interpersonal Problems. *British Journal of Clinical Psychology*, 35, 21–35. <https://doi.org/10.1111/j.2044-8260.1996.tb01159.x>
- Barlow, D. H., Nock, M. K., & Hersen, M. (2009). *Single case experimental designs. Strategies for studying behaviour change* (3rd ed.). Boston, MA: Pearson.
- Beck, A. T., Kovacs, M., & Weissman, A. (1979). Assessment of suicidal intention: The scale for suicide ideation. *Journal of Consulting and Clinical Psychology*, 47, 343–352. <https://doi.org/10.1037//0022-006x.47.2.343>
- Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. (1996). Comparison of Beck Depression Inventories-IA and-II in psychiatric outpatients. *Journal of Personality Assessment*, 67, 588–597. https://doi.org/10.1207/s15327752jpa6703_13
- Beck, A. T., Steer, R. A., & Brown, G. K. (1995). *Beck depression inventory II manual*. San Antonio, CA: The Psychological Corporation, Harcourt Brace & Co.
- Bennett, D., & Parry, G. (2004). A measure of psychotherapeutic competence derived from cognitive analytic therapy. *Psychotherapy Research*, 14, 176–192. <https://doi.org/10.1093/ptr/kph016>
- Berghoff, C. R., & Forsyth, J. P. (2015). Single-case designs: Methodology. In J. D. Wright (Ed.), *international encyclopedia of social and behavioral sciences* (2nd ed.; pp. 984–989). Oxford, UK: Elsevier.
- Black, D. W., Pfohl, B., Blum, N., McCormick, B., Allen, J., North, C. S., . . . Zimmerman, M. (2011). Attitudes toward borderline personality disorder: A survey of 706 mental health clinicians. *CNS Spectrum*, 16, 67–74. <https://doi.org/10.1017/S109285291200020X>

- Bloxham, G., Long, C. G., Alderman, N., & Hollin, C. R. (1993). The behavioral treatment of self-starvation and severe self-injury in a patient with borderline personality disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, 24, 261–267. [https://doi.org/10.1016/0005-7916\(93\)90030-z](https://doi.org/10.1016/0005-7916(93)90030-z)
- Bohus, M., Kleindienst, N., Limberger, M. F., Stieglitz, R. D., Domsalla, M., Chapman, A. L., . . . Wolf, M. (2009). The short version of the Borderline Symptom List (BSL-23): Development and initial data on psychometric properties. *Psychopathology*, 42, 32–39. <https://doi.org/10.1159/000173701>
- Borckardt, J. J., Nash, M. R., Murphy, M. D., Moore, M., Shaw, D., & O'Neil, P. (2008). Clinical practice as natural laboratory for psychotherapy research. *American Psychologist*, 63, 1–19. <https://doi.org/10.1037/0003-066X.63.2.77>
- Boschen, M., & Warner, J. (2009). Publication trends in individual DSM personality disorder: 1971–2015. *Australian Psychologist*, 44, 136–142. <https://doi.org/10.1080/00050060802680598>
- Bossart, D. F., Laird, V. C., & Armstrong, T. W. (2018). Interpreting Kendall's Tau and Tau-U for single-case experimental designs. *Cogent Psychology*, 5, 1–26. <https://doi.org/10.1080/23311908.2018.1518687>
- Brown, M. Z., Comtois, K. A., & Linehan, M. M. (2002). Reasons for suicide attempts and nonsuicidal self-injury in women with borderline personality disorder. *Journal of Abnormal Psychology*, 111, 198–202. <https://doi.org/10.1037//0021-843x.111.1.198>
- Calvert, R., & Kellett, S. (2014). Cognitive analytic therapy: A review of the outcome evidence base for treatment. *Psychology and Psychotherapy: Theory, Research and Practice*, 87, 253–277. <https://doi.org/10.1111/papt.12020>
- Chanen, A. M., Jackson, H. J., McCutcheon, L. K., Jovev, M., Dudgeon, P., Yuen, H. P., . . . Clarkson, V. (2008). Early intervention for adolescents with borderline personality disorder using cognitive analytic therapy: Randomised controlled trial. *The British Journal of Psychiatry*, 193, 477–484. <https://doi.org/10.1192/bjp.bp.107.048934>
- Christ, T. (2007). Experimental control and threats to internal validity of concurrent and nonconcurrent multiple baseline designs. *Psychology in the Schools*, 44, 451–459. <https://doi.org/10.1002/pits.20237>
- Cooper, M., Turpin, G., Bucks, R., & Kent, G. (2005). *Good practice guidelines for the conduct of psychological research within the NHS*. Leicester, UK: British Psychological Society, Research Ethics Practice Working Party of the Training Strategy Group.
- Davidson, K., & Scott, J. (2009). Does therapists' competence matter in delivering psychological therapy? *Psychiatric Bulletin*, 33, 121–123. <https://doi.org/10.1192/pb.bp.108.020214>
- Davidson, K. M., & Tyrer, P. (1996). Cognitive therapy for antisocial and borderline personality disorders: Single case study series. *British Journal of Clinical Psychology*, 35, 413–429. <https://doi.org/10.1111/j.2044-8260.1996.tb01195.x>
- Derogatis, L. R. (1993). *Brief Symptom Inventory: Administration, scoring and procedure manual* (4th ed.). Minneapolis, MN: National Computer Systems.
- Downey, G., Mougios, V., Ayduk, O., London, B. E., & Shoda, Y. (2004). Rejection sensitivity and the defensive motivational system: Insights from the startle response to rejection cues. *Psychological Science*, 15, 668–673. <https://doi.org/10.1111/j.0956-7976.2004.00738.x>
- Drucker, D. M. (2003). Testing for serial correlation in linear panel-data models. *Stata Journal*, 3, 168–177. <https://doi.org/10.1177/1536867X0300300206>
- Elliott, R., & Wexler, M. M. (1994). Measuring the impact of sessions in process experiential therapy of depression: The Session Impacts Scale. *Journal of Counseling Psychology*, 41, 166–174. <https://doi.org/10.1037/0022-0167.41.2.166>
- First, M. B., Gibbon, M., Spitzer, R. L., Williams, J. B. W., & Benjamin, L. S. (1997). *Structured clinical interview for DSM-IV axis II personality disorders, (SCID-II)*. Washington, DC: American Psychiatric Association.
- Fonagy, P., & Bateman, A. W. (2006). Mechanisms of change in mentalization-based treatment of BPD. *Journal of Clinical Psychology*, 62, 411–430. <https://doi.org/10.1002/jclp.20241>

- Heneghan, C., Goldacre, B., & Mahtani, K. R. (2017). Why clinical trial outcomes fail to translate into benefits for patients. *Trials*, 18, 122. <https://doi.org/10.1186/s13063-017-1870-2>
- Hersen, M. (1990). Single-case experimental designs. In A. S. Bellack, M. Hersen & A. E. Kazdin (Eds.), *International handbook of behavior modification and therapy* (pp. 175–210). Boston, MA: Springer.
- Hersen, M., & Barlow, D. H. (1976). *Single-case experimental designs: Strategies for studying behavior change*. New York, NY: Pergamon.
- Hughes, J., & Barkham, M. (2005). Scoping the inventory of interpersonal problems, its derivatives and short forms: 1988–2004. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice*, 12, 475–496. <https://doi.org/10.1002/cpp.466>
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12–19. <https://doi.org/10.1037//0022-006x.59.1.12>
- Jerschke, S., Meixner, K., Richter, H., & Bohus, M. (1998). The treatment history of patients with borderline personality disorder in the Republic of Germany. *Fortschritte der Neurologie-Psychiatrie*, 66, 545–552. <https://doi.org/10.1055/s-2007-995297>
- Jones, B., & Kenward, M. G. (2014). *Design and analysis of cross-over trials* (3rd ed.). London, UK: Chapman and Hall.
- Kazdin, A. E. (1978). Methodological and interpretive problems of single-case experimental designs. *Journal of Consulting and Clinical Psychology*, 46, 629–642. <https://doi.org/10.1037/0022-006X.46.4.629>
- Kellett, S. (2007). The treatment of histrionic personality disorder with cognitive analytic therapy; a time series evaluation. *Psychology and Psychotherapy: Theory, Research and Practice*, 80, 389–405. <https://doi.org/10.1348/147608306X161421>
- Kellett, S., & Beail, N. (1997). The treatment of chronic post-traumatic nightmares using psychodynamic-interpersonal psychotherapy: A single-case study. *British Journal of Medical Psychology*, 70, 35–49. <https://doi.org/10.1111/j.2044-8341.1997.tb01884.x>
- Kellett, S., Bennett, D., Ryle, T., & Thake, A. (2013). Cognitive analytic therapy for borderline personality disorder: Therapist competence and therapeutic effectiveness in routine practice. *Clinical Psychology & Psychotherapy*, 20, 216–225. <https://doi.org/10.1002/cpp.796>
- Kellett, S., & Hardy, G. (2014). The treatment of Paranoid Personality Disorder with CAT: A mixed methods single case experimental design. *Clinical Psychology and Psychotherapy*, 21, 452–64. <https://doi.org/10.1002/cpp.1845>
- Kellett, S., & Lees, S. (2019). Quasi-experimental N = 1 evaluation of the effectiveness of cognitive analytic therapy for dependent personality disorder. *Journal of Psychotherapy Integration*. Advance online publication. <https://doi.org/10.1037/int0000170>
- Kellett, S., Simmonds-Buckley, M., & Totterdell, P. (2016). Testing the effectiveness of cognitive analytic therapy for hypersexuality disorder: An intensive time series evaluation. *Journal of Marital and Sexual Therapy*, 43, 501–516. <https://doi.org/10.1002/cpp.1845>
- Kellett, S., & Totterdell, P. (2013). Taming the green-eyed monster: Temporal responsivity to cognitive behavioural and cognitive analytic therapy for morbid jealousy. *Psychology and Psychotherapy: Theory, Research and Practice*, 86, 52–69. <https://doi.org/10.1111/j.2044-8341.2011.02045.x>
- Kerr, I. B. (1999). Cognitive-analytic therapy for borderline personality disorder in the context of a community mental health team: Individual and organizational psychodynamic implications. *British Journal of Psychotherapy*, 15, 425–438. <https://doi.org/10.1111/j.1752-0118.1999.tb00473.x>
- Koehne, K., Hamilton, B., Sands, N., & Humphreys, C. (2013). Working around a contested diagnosis: Borderline personality disorder in adolescence. *Health*, 17(1), 37–56. <https://doi.org/10.1177/1363459312447253>
- Krasny-Pacini, A., & Evans, J. (2018). Single-case experimental designs to assess intervention effectiveness in rehabilitation: A practical guide. *Annals of Physical and Rehabilitation Medicine*, 61, 164–179. <https://doi.org/10.1016/j.rehab.2017.12.002>

- McMillan, D., & Morley, S. (2010). The quantitative single-case design as a research strategy for practice-based evidence. In M. Barkham, G. E. Hardy, & J. Mellor-Clark (Eds.), *Developing and delivering practice-based evidence: A guide for the psychological therapies* (pp. 109–138). Chichester, UK: Wiley-Blackwell.
- Mohr, D. C. (1995). Negative outcome in psychotherapy: A critical review. *Clinical Psychology: Science and Practice*, 2, 1–27. <https://doi.org/10.1111/j.1468-2850.1995.tb00022.x>
- Nehls, N. (2000). Recovering: A process of empowerment. *Advances in Nursing Science*, 22, 62–70. <https://doi.org/10.1097/00012272-200006000-00006>
- Nicklas, L. B., Dunbar, M., & Wild, M. (2010). Adherence to pharmacological treatment of non-malignant chronic pain: The role of illness perceptions and medication beliefs. *Psychology and Health*, 25, 601–615. <https://doi.org/10.1080/08870440902783610>
- Nordahl, H. M., & Nysæter, T. E. (2005). Schema therapy for patients with borderline personality disorder: A single case series. *Journal of Behavior Therapy and Experimental Psychiatry*, 36, 254–264. <https://doi.org/10.1016/j.jbtep.2005.05.007>
- Palihawadana, V., Broadbear, J. H., & Rao, S. (2019). Reviewing the clinical significance of ‘fear of abandonment’ in borderline personality disorder. *Australasian Psychiatry*, 27, 60–63. <https://doi.org/10.1177/1039856218810154>
- Paris, J., & Black, D. W. (2015). Borderline personality disorder and bipolar disorder. What is the difference and does it matter? *The Journal of Nervous and Mental Disease*, 203, 3–7. <https://doi.org/10.1097/NMD.0000000000000225>
- Pollock, P. H., Broadbent, M., Clarke, S., Dorrian, A., & Ryle, A. (2001). The Personality Structure Questionnaire (PSQ): A measure of the multiple self-states model of identity disturbance in cognitive analytic therapy. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice*, 8, 59–72. <https://doi.org/10.1002/cpp.250>
- Rizvi, S. L., & Linehan, M. M. (2005). The treatment of maladaptive shame in borderline personality disorder: A pilot study of “opposite action”. *Cognitive and Behavioral Practice*, 12, 437–447. [https://doi.org/10.1016/S1077-7229\(05\)80071-9](https://doi.org/10.1016/S1077-7229(05)80071-9)
- Rizvi, S. L., & Nock, M. K. (2008). Single-case experimental designs for the evaluation of treatments for self-injurious and suicidal behaviors. *Suicide and Life-Threatening Behavior*, 38, 498–510. <https://doi.org/10.1521/suli.2008.38.5.498>
- Robinson, L., Delgadillo, J., & Kellett, S. (2019). The dose-response effect in routinely delivered psychological therapies: A systematic review. *Psychotherapy Research*. Advance online publication. <https://doi.org/10.1080/10503307.2019.1566676>
- Rubel, J., Lutz, W., & Schulte, D. (2015). Patterns of change in different phases of outpatient psychotherapy: A stage-sequential pattern analysis of change in session reports. *Clinical Psychology and Psychotherapy*, 22, 1–14. <https://doi.org/10.1002/cpp.1868>
- Ryle, A. (1997). The structure and development of borderline personality disorder: A proposed model. *British Journal of Psychiatry*, 170, 82–87. <https://doi.org/10.1192/bjp.170.1.82>
- Ryle, A. (2007). Investigating the phenomenology of borderline personality disorder with the states description procedure: Clinical implications. *Clinical Psychology and Psychotherapy*, 14, 329–341. <https://doi.org/10.1002/cpp.543>
- Ryle, A., Beard, H., & Marlowe, M. J. (1995). Cognitive analytic therapy of borderline personality disorder: Theory, practice and the clinical and research uses of the self-states sequential diagram. *International Journal of Short-Term Psychotherapy*, 10, 21–34.
- Ryle, A., & Golyunkina, K. (2000). Effectiveness of time-limited cognitive analytic therapy of borderline personality disorder: Factors associated with outcome. *British Journal of Medical Psychology*, 73, 197–210. <https://doi.org/10.1348/000711200160426>
- Ryle, A., & Kellett, S. (2018). Cognitive analytic therapy. In W. John Livesley & R. Larstone (Eds.), *Handbook of personality disorders: Theory, research, and treatment* (2nd ed., pp. 489–512). New York, NY: Guilford Press.
- Ryle, A., Kellett, S., Hepple, J., & Calvert, R. (2014). Cognitive analytic therapy at 30. *Advances in Psychiatric Treatment*, 20, 258–268. <https://doi.org/10.1192/apt.bp.113.011817>

- Ryle, A., & Kerr, I. B. (2003). Cognitive analytic therapy. *British Journal of Psychiatry*, 183, 79. <https://doi.org/10.1192/bjp.183.1.79>
- Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is it really robust? Reinvestigating the robustness of ANOVA against violations of the normal distribution assumption. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 6, 147–151. <https://doi.org/10.1027/1614-2241/a000016>
- Shadish, W. R., & Sullivan, K. J. (2011). Characteristics of single-case designs used to assess intervention effects in 2008. *Behavior Research Methods*, 43, 971–980. <https://doi.org/10.3758/s13428-011-0111-y>
- Shaw, C., & Proctor, G. (2005). Women at the margins: A critique of the diagnosis of borderline personality disorder. *Feminism & Psychology*, 15, 483–490. <https://doi.org/10.1177/0959-353505057620>
- Skodol, A. E., Gunderson, J. G., Pfohl, B., Widiger, T. A., Livesley, W. J., & Siever, L. J. (2002). The borderline diagnosis I: Psychopathology, comorbidity, and personality structure. *Biological Psychiatry*, 51, 936–950. [https://doi.org/10.1016/s0006-3223\(02\)01324-0](https://doi.org/10.1016/s0006-3223(02)01324-0)
- Spence, C., Kellett, S., Totterdell, P., & Parry, G. (2019). Can cognitive analytic therapy treat Hoarding Disorder? An adjudicated hermeneutic single-case efficacy design evaluation. *Clinical Psychology & Psychotherapy*, 26, 673–683. <https://doi.org/10.1002/cpp.2390>
- Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Barlow, D. H., . . . Wilson, B. (2016). The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016 statement. *Archives of Scientific Psychology*, 4, 1–9. <https://doi.org/10.1037/arc0000026>
- Widiger, T. A. (2015). Assessment of DSM–5 personality disorder. *Journal of Personality Assessment*, 97, 456–466. <https://doi.org/10.1080/00223891.2015.1041142>
- Wildgoose, A., Clarke, S., & Waller, G. (2001). Treating personality fragmentation and dissociation in borderline personality disorder: A pilot study of the impact of cognitive analytic therapy. *British Journal of Medical Psychology*, 74, 47–55. <https://doi.org/10.1348/000711201160795>
- Wirth-Cauchon, J. (2001). *Women and borderline personality disorder*. New Brunswick, NJ: Rutgers University Press.

Received 30 October 2019; revised version received 16 March 2020

Supporting Information

The following supporting information may be found in the online edition of the article:

Figure S1. Partial autocorrelation function (PACFs) for idiographic measures.

Appendix S1. Session impact scores and comments.