# Accepted Manuscript


Time-intensive behavioural activation for depression: A multiple baseline study

S. Miles, G. Brown, A. Corfe, C. Hallett, J. Wingrove, J. Wheatley, D. Veale

PII: S0005-7916(18)30156-3

DOI: https://doi.org/10.1016/j.jbtep.2018.12.008

Reference: BTEP 1447

To appear in: *Journal of Behavior Therapy and Experimental Psychiatry*

Received Date: 9 June 2018

Revised Date: 11 November 2018

Accepted Date: 22 December 2018

Please cite this article as: Miles, S., Brown, G., Corfe, A., Hallett, C., Wingrove, J., Wheatley, J., Veale, D., Time-intensive behavioural activation for depression: A multiple baseline study, *Journal of Behavior Therapy and Experimental Psychiatry* (2019), doi: https://doi.org/10.1016/j.jbtep.2018.12.008.

# Time-intensive behavioural activation for depression: A multiple baseline study.

Miles, S.[a], Brown, G.[a], Corfe, A.[b], Hallett, C.[c], Wingrove, J.[c], Wheatley, J.[b], & Veale, D.[c & d]

[a] Royal Holloway University of London, Egham, Surrey, United Kingdom

[b] Homerton University Hospital NHS Foundation Trust, London, United Kingdom

[c] South London and Maudsley NHS Foundation Trust, London, United Kingdom

[d] Institute of Psychiatry, Psychology and Neurosciences, London, United Kingdom

Address for correspondence: Sarah Miles, Royal Holloway University of London, University Campus, Egham, Surrey, TW20 0EX

Email: sarah.miles.2014@live.rhul.ac.uk

# Time-intensive behavioural activation for depression: A multiple baseline study.

Abstract

Background and Objectives: Depression is the second leading cause of disability, worldwide, and increasing access to its effective/preferred treatment requires more attention. Behavioural activation and time-intensive treatment delivery both show promise in this regard, yet research into their combination is limited. This study aimed to investigate the feasibility, effectiveness, and acceptability of time-intensive behavioural activation (BA) for depression. Methods: Eight adults with major depressive disorder were recruited from three outpatient IAPT services in London. The study employed a single case experimental design with multiple baselines. All participants completed time-intensive BA, consisting of up to seven twice weekly sessions with daily prompting in-between and three optional booster sessions. Idiographic, standardised and process measures of depression symptomatology were collected. Results: Treatment recruitment and retention indicated that the intervention was feasible. Visual and statistical analyses showed that relative to baseline, 6 out of 8 participants made significant improvements in all idiographic symptoms of depression following the intervention. According to standardised measures of depression, four out of eight participants were considered treatment responders. Five participants completed follow-up measures and the majority of progress was maintained after the withdrawal of the intervention. The intervention was also considered highly acceptable by participants and therapists. Limitations: Conclusions cannot be drawn about the generalizability or the long-term durability of the findings. Conclusions: Overall this study provides new, but tentative evidence highlighting the potential of time-intensive BA as a feasible, effective and acceptable treatment for *some* adult outpatients with depression. The findings now warrant further, more rigorous evaluation of the treatment.

Keywords: depression; behavioural activation; BA, intensive; single case design; multiple baseline.

## 1. Introduction

Depression is the second leading cause of disability worldwide (Ferrari et al., 2013), and its ubiquitous burden is costly (Bloom et al., 2011). Despite the availability of a range of evidence-based treatments for depression (NICE, 2009) and the Increasing Access to Psychological Therapies (IAPT) initiative (DOH, 2008) in England the efficacy of psychological treatments for depression appear to have reached a plateau (Cuijpers, 2015). Response rates following depression treatments within IAPT are 55% (Richards & Borglin, 2011), and, on average, 17.5% of depressed clients drop out of treatment (Cooper & Conklin, 2015). Substantial numbers of depression cases remain undetected (Glover, Webb, & Evison, 2010), and many practical barriers to effective treatment remain, including characteristics of depression itself (e.g., pessimism), time constraints, and personal responsibilities (e.g., child care and work schedules) (Mohr et al., 2010).

While antidepressant medication (ADM) and cognitive behavioural therapy (CBT) have demonstrated the most convincing empirical support to date (NICE, 2009), behavioural activation has been suggested to hold particular promise as an accessible and disseminable treatment for depression (Kanter, Puspitasari, Santos, & Nagy, 2012). Indeed, its most common variants, Martell, Addis and Jacobson's (2001) behavioural activation, (herein referred to as BA), and the briefer behavioural activation treatment for depression (herein referred to as BATD; Lejuez, Hopko, LePage, Hopko & McNeil, 2001) are found to be effective when disseminated across diverse settings populations and treatment formats (Dimidjian, Barrera, Martell, Munoz & Lewinsohn, 2011).

[1]The abbreviations BA and BATD will be used to refer to specific versions of behavioural activation. The unabbreviated term will be used for the general category of approaches.

Dobson et al. (2008) found that BA had lower drop-out rates and more durable effects than ADM, and a recent randomised controlled trial (RCT) found that BA produced comparable results to CBT even if delivered by less highly trained professionals than CBT requires, and that at 12 months post-treatment it is more cost-effective than CBT (Richards et al., 2016).

Increasing treatment intensity is a potential means for enhancing treatment efficacy that is increasingly being explored. Regression analyses have demonstrated that faster overall recovery is significantly associated with more frequent psychological treatment delivery (Erekson, Lambert, & Eggett, 2015). Time-intensive psychological interventions, operationalized here as treatment delivered more frequently and concentrated over a shorter period of time than the traditional weekly 50-minute session, have demonstrated efficacy in the treatment of PTSD (Ehlers et al., 2014), OCD (Storch et al., 2007) and specific phobias (Zlomke & Davis, 2008). Number, frequency, and duration of sessions have varied widely, making it difficult to draw general conclusions regarding how best to structure intensive treatment. However, because depression and anxiety treatments include key elements that are functionally similar (see Hopko, Lejuez, Ruggiero, & Eifert, 2003b for a comparison of behavioural activation and exposure) it is reasonable to assume that time-intensive BA could have positive effects for some individuals. BA itself is already recommended to be delivered twice a week for the first three to four weeks, and then once a week thereafter (NICE, 2009). However, this recommendation is not followed consistently in clinical practice (Dimidjian, Martell, Addis, Herman-Dunn, & Barlow, 2008).

Time-intensive treatment for depression has not been investigated to the same extent as it has been for anxiety. One meta-regression analysis concluded that effectiveness of depression treatment was more strongly related to session *intensity* (number of sessions a week) than general treatment *quantity* (Cuijpers, Huibers, Ebert, Koole & Andersson, 2013); delivering two sessions rather than one a week was associated with increased treatment effect size ($g$ = 0.45). However, only six studies included in this analysis had delivered behavioural activation. In addition, the findings were based on planned treatment parameters rather than the characteristics of the treatment actually provided and so did not provide a clear-cut basis for associating particular therapy parameters (number, frequency and duration of sessions) with different outcomes. Cuijpers et al., 2013 concluded that the potential gain from optimal pacing, and particularly intensifying, of existing therapies is a more auspicious target of research than developing and testing new modalities for depression, although they note that there is a likely lower limit such that "no one would probably consider treating depression in one week" (p. 11). However, optimizing intensity coupled with greater specificity of knowledge within BA about the temporal sequencing of activation relative to symptom change (e.g., Santos et al., 2017) offers the prospect of maximizing efficient deployment of service delivery.

There is some evidence, though with limitations, for the efficacy of a single 90-minute BATD session relative to no treatment (Gawrysiak, Nicholas, & Hopko, 2009) and waitlist controls (Nasrin, Rimes, Reinecke, Rinck, & Barnhofer, 2017). One study that did not find support for the single session significantly reducing depression symptoms in comparison to a waitlist control, proposed that future research should include short telephone calls between the therapist and client,

to prompt activation (Read, Mazzucchelli, & Kane, 2016).

Time-intensive behavioural activation has been more commonly researched in inpatient settings, where, as length of stay is reduced, session frequency tends to be higher. One RCT demonstrated BATD's efficacy, in comparison to a supportive psychotherapy control condition, when delivered to 25 participants via 20-minute sessions, three times a week, over a two-week period (Hopko, Lejuez, Lepage, Hopko, & McNeil, 2003a). However, there was a lack of an empirically validated control condition, and treatment adherence was not assessed. More recently, a rigorous multiple baseline single case experimental design (SCED) investigated the efficacy of a time-intensive behavioural treatment for depression based on a synthesis of BA and BATD (Folke et al., 2015). Six inpatients with depression were randomised to different baseline lengths followed by daily 20 - 30 min therapy sessions over five consecutive days. Five of the six showed reliable change in self-reported and depression symptomatology and four of six on clinicians' ratings, with one fewer demonstrating *clinically significant* change on each measure. Participants rated the treatment as highly satisfying. However, there was potential bias stemming from non-random recruitment of participants and delivery of therapy by the first author of the study. There are also several general limitations evident in this group of studies, including their small sample sizes and lack of follow-up data to examine maintenance of gains. Further, none of the studies used structured clinical interviews to diagnose MDD, and participants tended to have comorbidities and were receiving multiple inpatient treatments concurrent with their behavioural activation, potentially confounding the findings.

Despite more rigorous existing evidence of *BA's* efficacy (Richards et al., 2016), particularly within adult outpatient settings (Kanter et al., 2010), and most treatment of depression occurring therein, this review found no examples of studies investigating the effects of time-intensive BA in outpatient settings. Given the implication that time-intensive BA for depression could enhance treatment access and effects; this gap in the literature is worth addressing.

Therefore, the aims of this study were to explore whether or not time-intensive BA was (a) a feasible intervention, (b) had an effect on idiographic measures of depression symptoms, (c) was associated with reliable and clinically significant change in standardised and process measures of depression, (d) maintained its effects over a follow-up period, and (e) was an acceptable treatment.

A multiple baseline single case experimental design (SCED) was used. SCEDs entail monitoring change within participants through repeated measurement of outcomes across different study phases (Kazdin, 1982), but also capture individual differences between participants (Morley, 2015a). They are a robust method for testing causal mechanisms of treatments as they enable participants to act as their own controls. Randomising participants to multiple baselines increases experimental control over potential extraneous threats to validity, such as maturation (Kratochwill et al., 2010).

It was hypothesised that following the intervention, participants would show improvements in their idiographic, standardised and process measures of depression symptomatology. Improvement was operationalised as making significant declines on idiographic measure ratings (according to *Tau* statistics), or reliable change (Jacobson & Traux, 1991) on standardised and process measures.

## 2. Material and Methods

### 2.1 Participants

Participants seeking treatment for depression were recruited from three London IAPT services. Initially only two services were approached, but later a third site was added. Sixty consecutive clients presenting for treatment of depression, within the recruitment period, were offered the opportunity to take part in the study as a treatment option. Twenty-three of these clients were interested in taking part in the study, and were screened. Only nine of these clients met the inclusion criteria and no exclusion criteria, and so were invited to be participants, eight of whom then consented to take part (Figure 1). The one eligible participant who declined taking part could no longer attend sessions due to returning to work. All participants were either unemployed or still in education (Table 1).

The study received ethical approval from the UK Central London Research Ethics Committee.

#### 2.1.1 Inclusion criteria

(1) Age 18 or over;

(2) Primary diagnosis of MDD as assessed by the Structured Clinical Interview for MDD (First, Williams, Karg, & Spitzer, 2015);

(3) Sufficient command of English to comprehend instructions and measures;

(4) A score of ≥ 10 on the Patient Health Questionnaire (PHQ-9; Kroenke & Spitzer, 2002);

(5) A score of ≥ 25 on the MADRS (Montgomery & Åsberg, 1979);

(6) No psychopharmacological treatment or on a stable dose for six weeks prior to the study, with the type and dosage unaltered;

(7) Ability to travel to treatment at the time-intensive rate.

2.1.2 Exclusion criteria

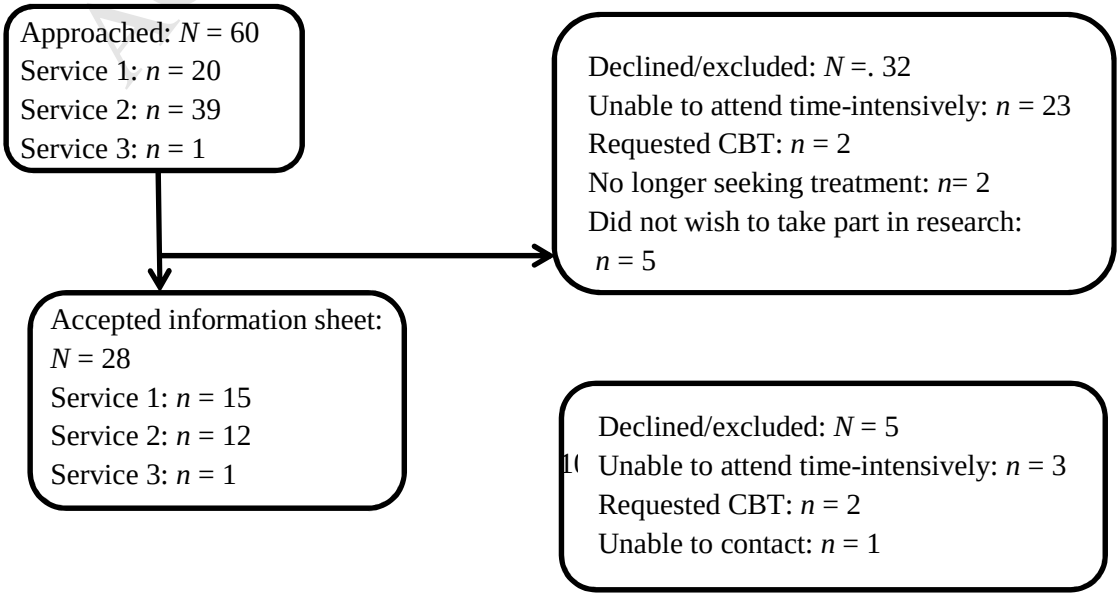
(1) Major suspected comorbid diagnosis of severe disorder (e.g., personality disorder or substance dependency) according to screening (which included use of the Standardised Assessment of Personality – Abbreviated Scale [SAPAS; Moran, Leese, Lee, Walters, Thornicraft, & Mann, 2003] and clinical judgement;

(2) Acutely suicidal;

(3) Concurrent additional psychotherapy;

(4) Diagnosis of a long-term condition that would prevent attendance (e.g., requiring hospitalisation).

Participants were not offered any compensation for taking part in the study.



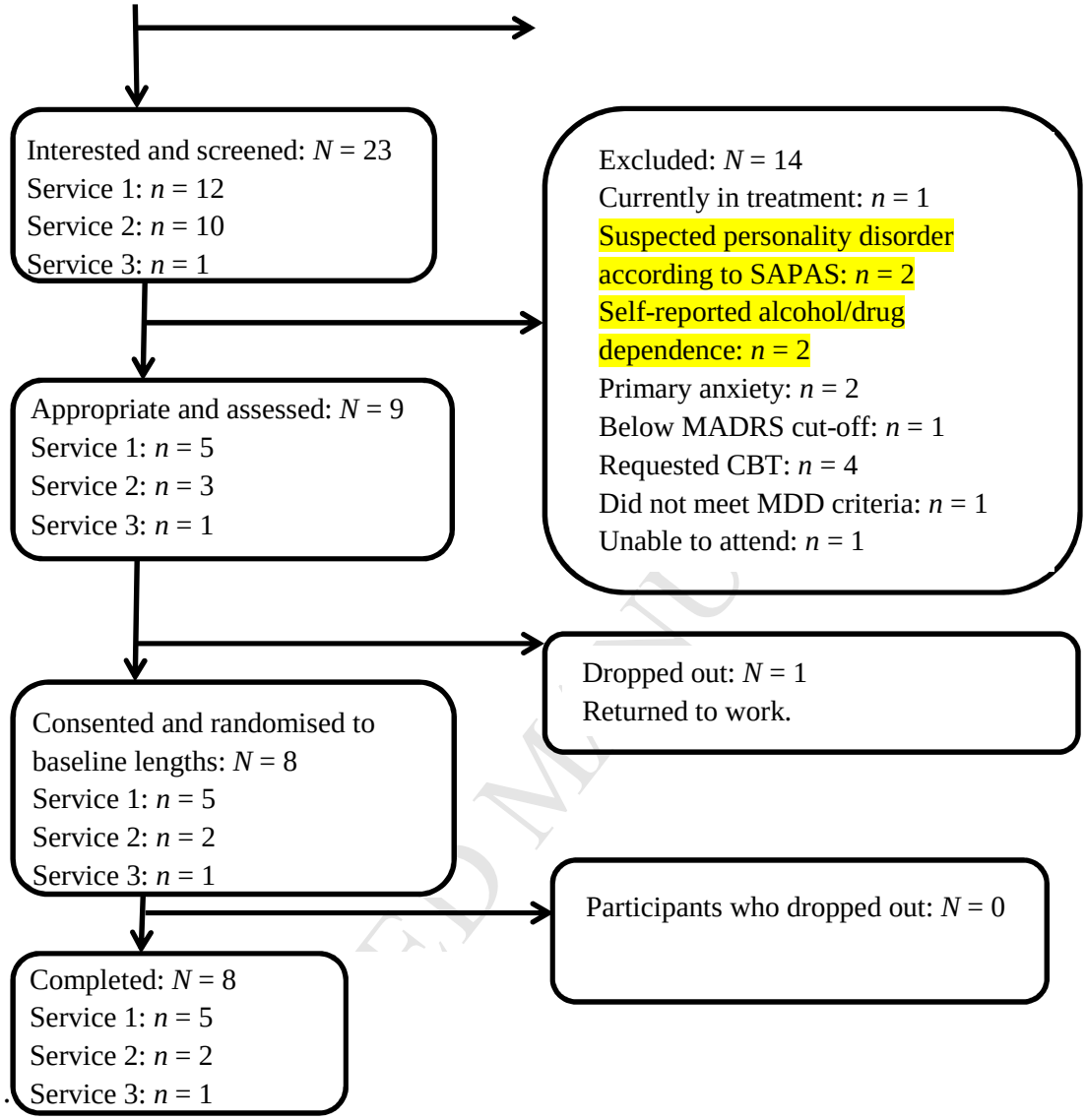



*Figure 1.* Flow of recruitment and participation

Table 1 *Clinical summary of participants*

| P | Service | Sex | Age | Ethnicity | Marital status | Education Level | Employment status | Duration of problem | Previous episodes | Previous treatment (year) | Current medication (dose, duration) | Comorbidity |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A (PA) | 1 | F | 37 | White Portuguese | Single | Undergraduate degree | Unemployed | 2 months | 2 | 1.CBT for anxiety (2015) | None | GAD<br>LTC |
| B (PB) | 1 | F | 60 | White British | Divorced | O-Levels | Retired | 10 months | 3 | 1. Psychiatric hospitalisation for depression (1981)<br>2. Counselling (1990) | Citalopram (20 mg, 12 months) | LTC |
| C (PC) | 2 | F | 49 | White Mixed European | Separated | Undergraduate degree | Unemployed | 4 years | 4 | 1.CBT for low mood (2015) | Citalopram (20 mg, 9 months) | GAD<br>LTC |
| D (PD) | 2 | F | 31 | Black British / Caribbean | In a Relationship | NVQ | Long-term Sick Leave | 4 months | 4 | 1.CBT for low mood (2011) | Fluoxetine (20 mg, 10 weeks) | Secondary anxiety and panic attacks |
| E (PE) | 1 | M | 21 | White British | Single | A-levels | Unemployed | 2 months | 3 | 1.CBT for depression (2014) | Citalopram (10 mg, 8 weeks) | None |
| F (PF) | 1 | F | 27 | White European | Single | Undergraduate degree | Student | 6 months | 2 | None | None | Secondary anxiety |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| G (PG) | 3 | F | 28 | White British | Cohabiting with Partner | Postgraduate degree | Student | 1 year | 3 | 1.Counselling for bereavement (2014) | None | GAD |
| H (PH) | 1 | M | 56 | White British | In a Relationship | A-levels | Unemployed | 6 months | 2 | 1.Counselling (1992)<br>2. Counselling (2016) | Fluoxetine (40 mg, 5 months) | GAD<br>LTC |
| Mean (*SD*) | | | 39 (14.57) | | | | | 11 months (14.27) | 3 (0.83) | | | |

*Note. GAD = generalized anxiety disorder; LTC = long-term physical health condition; NVQ = National Vocational Qualification; P = participant; SD = standard deviation. Some details have been changed, to protect participant anonymity.*

## 2.2 Design

The study employed an $A_1BA_2$ SCED. The $A_1$ phase was the multiple baseline phase with participants randomly allocated to baseline lengths of seven, 14 or 21 days. Phase B was the intervention phase (time-intensive BA) followed by a three-week follow-up period of symptom monitoring ($A_{2)}$.

### 2.2.1 Therapists

Six, out of 23 therapists approached, were involved in the study. Reasons for not taking part in the study were working part-time ($n$ = 4), having a full case-load ($n$ =10), and not yet being fully qualified ($n$ = 3). Inclusion criteria included that all therapists were either qualified Clinical Psychologists or High-Intensity Cognitive Behavioural Therapists. They had received training in BA as part of their degrees but were also required to attend a half-day BA training session.

## 2.3. Measures

Clinician rated, and self-report measures of depression were collected.

### 2.3.1. Idiographic measures

The primary outcome measures were visual analogue scales (VASs) of depression, rumination, avoidance, a participant chosen main symptom of depression, and an encapsulated belief. Encapsulated beliefs consisted of statements summarising the meaning of participants' experiences of depression. Scores ranged from 0 to 100, with higher scores indicating higher symptomatology or frequency of belief. VASs were completed daily across all three study phases.

### 2.3.2 Standardized measures

Standardized measures of depression severity were completed weekly throughout the baseline, intervention and follow-up phases of the study, to provide global context to the interpretation of idiographic measure outcomes. These were:

***(a) The Montgomery-Åsberg Depression Rating Scale (MADRS; Montgomery & Åsberg, 1979).*** The MADRS is a 10-item clinician-rated scale measuring depression symptoms over the past week. Each item is rated on a seven-point Likert scale from 0 (indicating *'normal'* or *'no difficulties'*) to 6. The summed score range is 0–60, and higher scores reflect greater symptomatology. The measure has shown high inter-rater reliability between .89 and .97, and significant convergent validity (Montgomery & Åsberg, 1979).

***(b) The Patient Health Questionnaire (PHQ-9; Kroenke & Spitzer, 2002)***. The PHQ-9 is a 9-item self-report measure of depression symptoms, usually from the past two weeks, yet here participants were asked to refer to the past week. Each item is scored from 0 ('*not at all'*) to 3 ('*nearly every day'*), and summed scores range from 0 to 27, with higher scores reflecting greater symptomatology. The measure has high reported internal consistency (Cronbach's $\alpha = .89$) and convergent validity (Kroenke, Spitzer, & Williams, 2001).

2.3.3 Process measures

***The Behavioural Activation for Depression Scale - Short Form (BADS-SF; Manos, Kanter & Luo, 2011).*** The BADS-SF is a 9-item self-report scale measuring activation and avoidance over the past week used to determine whether BA is having an impact on activity levels. It was completed weekly over the baseline and intervention phases. Items are rated from 0 ("*not at all*")

to 6 ("*completely*"). Total scores range from 0 to 54. Higher scores represent more activation. It has good internal consistency (Cronbach's $\alpha$ = .82) and demonstrated construct validity (Manos et al., 2011).

2.3.4 Ending measures

The following measures were completed after participants finished their final session:

***(a) The Client Satisfaction Questionnaire (CSQ; Larsen, Attkisson, Hargreaves, & Nguyen., 1979).*** The CSQ is an 8-item self-report measure of clients' satisfaction with services. Items are rated on different four-point Likert scales and are summed to give a total score ranging from 8 to 32, with higher scores indicating greater satisfaction. Reported Cronbach's α ranges from .83 to .93, indicating excellent internal consistency.

***(b)*** Participants and therapists were also asked to rate how acceptable they found the intervention, on VASs, ranging from 0 ("*not at all*") to 100 ("*completely*").

***(c) The Quality of Behavioural Activation Scale (QoBAS; Dimidjian, Hubley, Martell & Herman, 2016).*** The QoBAS is a 14-item scale assessing therapists' use of BA strategies. Items are rated on a seven-point Likert Scale from 0 ("*poor*") to 6 ("*excellent*"), with a score of 3 indicating satisfactory BA skill quality. Higher scores indicate greater treatment quality. The measure is not yet validated but is currently the primary measure of BA fidelity in use.

2.4. Intervention

Time-intensive BA for depression was delivered, based on Martell et al.'s (2013) clinicians' guide. BA entails carrying out idiographic functional analyses of participants' depressive behaviours (including cognitive processes such as

rumination), as well as the contextual contingencies of reinforcement maintaining them. It then promotes engagement in activities and contexts that promote behaviours that counteract depression and that are consistent with the client's values and goals. Activity scheduling is regularly restructured according to individual formulations to promote completion of scheduled activities.

Three independent assessors used the QoBAS (Dimidjian et al., 2016) to assess one randomly selected session from five participants (meaning that 7% of all sessions were assessed). The *mean* QoBAS score was 3.74 (*SD* = 0.75), exceeding the satisfactory treatment competence threshold. In addition, 80% of assessed sessions (*n* = 4) met minimum criteria of a rating of 3.00. Reliability between the assessor's ratings was also acceptable, with an *intra-class correlation* of .83.

2.5. Procedure

Clients seeking treatment for depression identified as suitable for BA, were sent information sheets. Those willing and eligible to participate attended a study assessment and were randomly allocated to an intervention start time. The therapist and participant then worked collaboratively to build a functional analysis. Participants identified and completed their idiographic measures and were given a timetable of their scheduled sessions and outcome measures to be completed.

Therapists aimed to deliver a minimum of 10 and maximum of 13 hours of face-to-face BA. The initial seven sessions were twice weekly, to allow ample activity monitoring and activation implementation, with the initial three sessions recommended to last for two hours, and the last four recommended to last one hour. The inclusion of a break in longer sessions was suggested to counteract fatigue. Three

optional booster sessions were offered and recommended to take place at one, two and four weeks after a participant's seventh session.

On each weekday, in-between face-to-face sessions, and up until booster sessions began, therapists or an Assistant Psychologist (see Supplementary Table 1) telephoned, texted, or emailed participants for short 'prompting' conversations, whenever they could. Prompting was intended to as a method of contingency management, to demonstrate therapists' dedication to participants, to give participants direction, and to positively reinforce progress, morale, motivation, treatment compliance, and the therapeutic relationship. This is in line with one of the core principles of BA (Martell et al., 2010), which is for therapists to act as a "coach". It also meant that participants who had limited support networks still had a regular form of reinforcement outside of session times.

The frequency of prompts was not pre-defined, and was dependent on need and feasibility. However, therapists logged prompting durations and aimed to keep each prompting telephone conversation to ten minutes long. Participants could respond to text messages or emails whenever they wanted to but understood that therapists would attempt to respond within 24 hours.

In line with previous research, participants who received eight hours of clinician contact were deemed treatment completers (Richards et al., 2016). Where possible, symptom monitoring continued for three weeks post treatment, and all participants were offered a clinical review after follow-up to assess need for further treatment. If so, participants could continue sessions with their existing therapist or necessary onward referrals were made.

## 2.6 Analysis

Idiographic data were first graphed using Microsoft Excel, and then visual analyses were conducted to determine whether or not the intervention was effective. Change in the central tendency (reflecting the level of symptom change), trend (reflecting the strength of symptom change) and variability (reflecting the stability of symptom change) of idiographic symptoms were assessed, across each phase, in accordance with visual analysis guidelines (Kazdin, 1998). Statistical analysis of idiographic data was conducted using Tau-U (Parker, Vannest, Davis, & Sauber, 2011), which tests for the percentage of non-overlap between study phases, controlling for baseline trend. Weighted averages, forming omnibus effect sizes across all of the participants, were also calculated for each idiographic measure (http://www.singlecaseresearch.org/calculators/tau-u).

The MADRS, PHQ-9 and BADS-SF were used to identify the number of participants who displayed reliable and clinically significant change (Jacobson & Traux, 1991) from baseline to end of treatment and follow-up. Comparisons of pre and post-intervention scores were based on single or average baseline scores (where participants had two). Post-BA and follow-up scores were calculated from final session and final follow-up session scores, respectively. Criterion “a” was used for reaching reliable and clinically significant change on the MADRS and BADS-SF, whereas criterion “b” was used for the PHQ-9 (Jacobson & Traux, 1991). Clinical and non-clinical population means and standard deviations, from existing research data, were used to calculate the criteria (Cunningham, Wernroth, von Knorring, Berglund, & Ekselius, 2011; Dimidjian et al., 2017; McMillan, Gilbody, & Richards, 2010).

Participants were classified as treatment responders if they met criteria for both reliable and clinically significant change on one or both measures of depression.

## 3. Results

### 3.1. Treatment feasibility

Table 1 provides a brief clinical summary of each participant.

All participants were considered treatment completers. In total, 68 of 76 sessions offered were attended (89%). Variation in attendance and prompting frequencies resulted in variations in treatment durations across participants (Supplementary Table 1). On average, participants received nine sessions (*SD* = 1.41), or 11 hours and 20 minutes of face-to-face therapist contact (*SD* = 1.62). All participants received different types and frequencies of prompting (see Supplementary Table 1). However, an average of 17 prompting text messages were sent to participants being prompted via text alone, and an average of six prompting telephone calls were made to participants being prompted over the telephone. Due to limited data, it was not possible to calculate average prompting frequencies for participants receiving both text and telephone call prompts, or email prompts.

### 3.2. Idiographic symptom outcomes

Figures 2 to 5 display raw data, lines of central tendency and trend for participants' idiographic depression and rumination ratings. All eight participants showed some fluctuation in idiographic depression ratings over the baseline phase, but participants A, B, E, and G showed slight upward trends. During the intervention phase, all participants except participant C (whose symptoms remained relatively

unchanged), showed declines in trend and central tendencies, relative to baseline, and statistical analyses indicated that they experienced significant reductions in depression (Table 3). Declines in trend were not consistent. The change in trend was more noticeably pronounced and rapid for participant B. Participant E showed a more stepped decline after losing a friend to suicide. Participants F, G and H showed more gradual declines, and participant A's trend was less clear. Four participants who provided follow-up data did not show significant changes in their depression scores following the withdrawal of the intervention. Conversely, Participant D's scores showed reversed upward trend over the last two weeks of her treatment, and this continued to increase at follow-up.

All eight participants also showed fluctuation in their rumination ratings over the baseline phase, and participants B, E, F and G's rumination showed slight increasing trends. During the intervention phase, all participants showed declines in levels of rumination, relative to the baseline phase; however, these only reached significance for six participants (Table 4). Patterns of change seen for depression ratings were reflected in rumination ratings; however participant C showed small declines in levels of rumination over the intervention phase, which increased again at follow-up, and participant F experienced further significant reduction in her rumination following the withdrawal of the intervention.

VAS's of avoidance, encapsulated beliefs, and chosen symptoms of depression also showed that the majority of participants ($n$ = 5, 6, and 7, respectively) experienced significant reductions in symptoms, following the onset of the intervention, and that the majority of participants who provided follow-up data reported maintained or continued improvements (Supplementary Tables 2 to 4).

Weighted averages of non-overlap of data between phases (Tables 3 to 4 and Supplementary Tables 2 to 4) indicated that, across all participants, all idiographic symptoms showed significant declines between baseline and intervention phases, with depression and encapsulated beliefs decreasing the most. Between the intervention and follow-up phases, all symptom ratings reflected decline in the intervention's effects, but the proportion of depression ratings that increased reached significance. When combining data from intervention and follow-up phases, relative to baseline, significant declines were still evident on all idiographic symptom ratings.

Of note, all participants reported experiencing ongoing life-stressors, which may have influenced the high variability of data, and biased weighted average calculation. For example, PD's pattern of symptom change may have been influenced by her experience of a traumatic event at approximately day 49. As the only participant to show significant deterioration in depression at follow-up, her outcomes may have influenced the significant increase in the weighted average of non-overlap of depression scores between intervention and follow-up. Therefore, individual outcomes, as opposed to weighted averages were considered for the basis of drawing conclusions (Parker & Vannest, 2012).

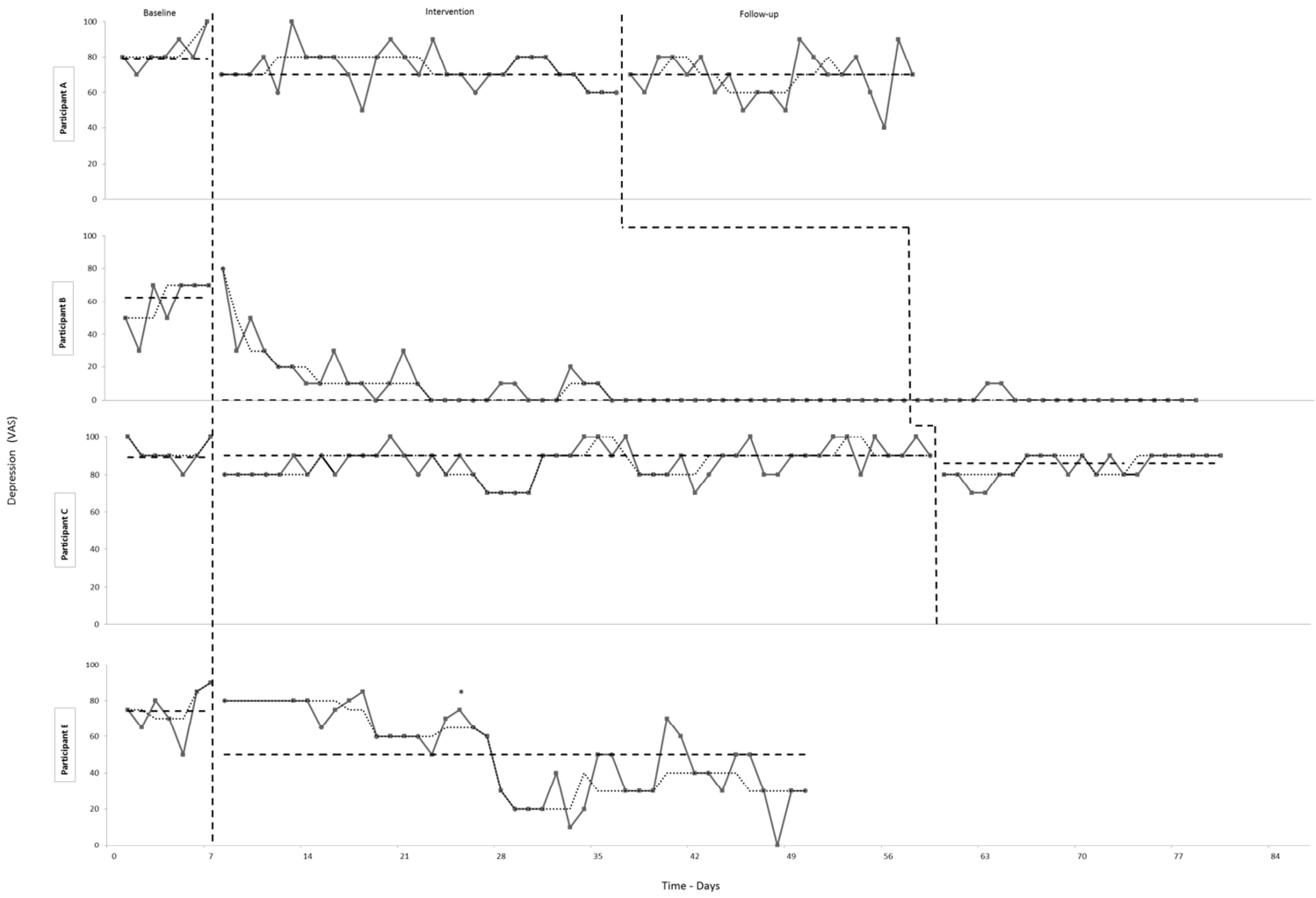



*Figure 2.* Depression VAS data from participants A, B, C and E.

*Note.* Raw data = ; central tendency (broadened median) = ; trend (running median) = ; = session; * = life stressor. Participant E lost a friend to suicide and reported that it made him feel less abnormal. Different calculations of central tendency and trend plots were chosen according to Morley's visual analysis guidelines (Morley, 2015b).

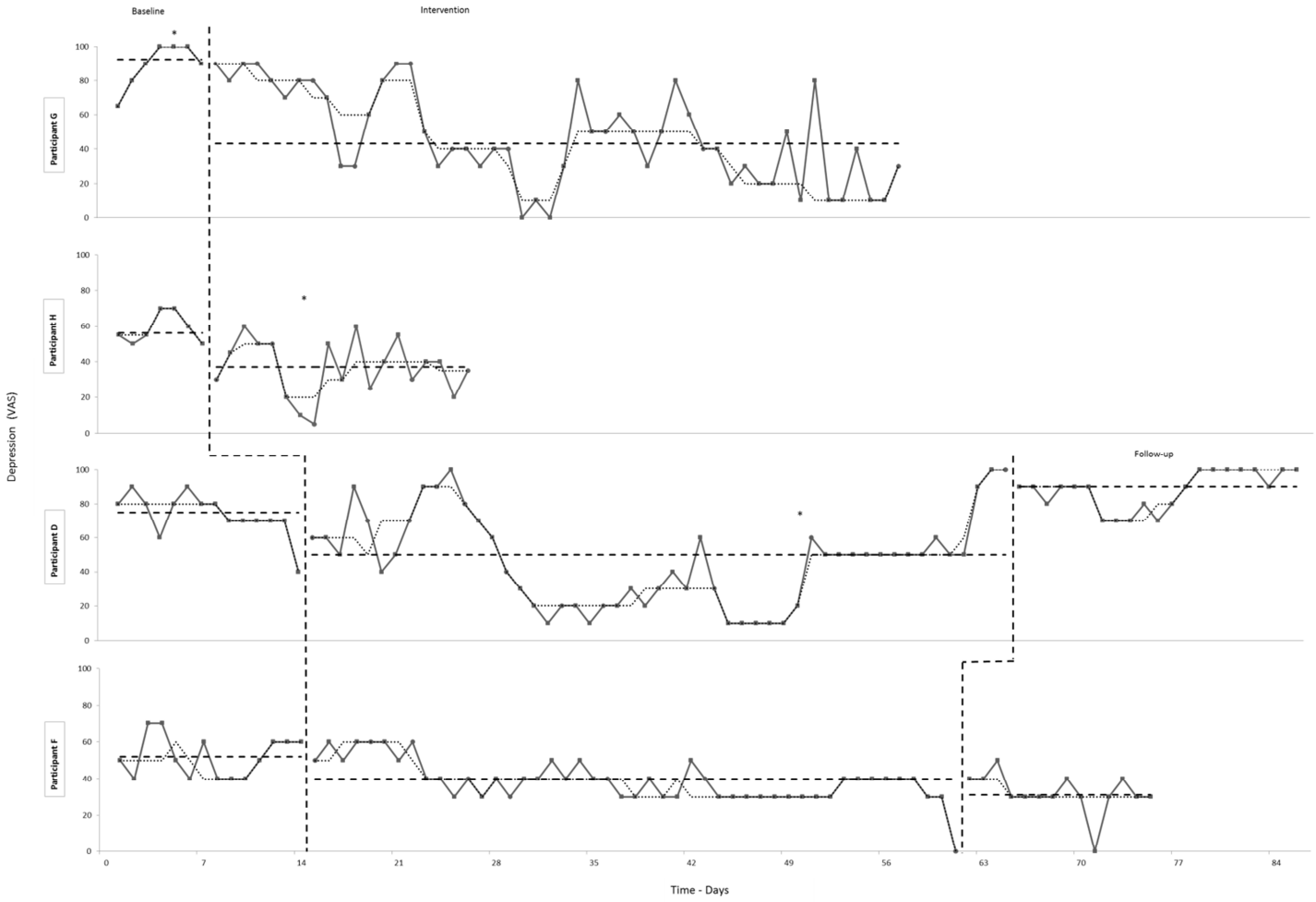



*Figure 3*. Depression VAS data from participants G, H, D and F.

*Note.* Raw data = ; central tendency (broadened median) = ; trend (running median) = ; = session; * = life stressor. Participant D experienced a trauma, Participant G experienced a relationship breakdown and participant H stopped taking his medication. Different calculations of central tendency and trend plots were chosen according to Morley's visual analysis guidelines (Morley, 2015b).

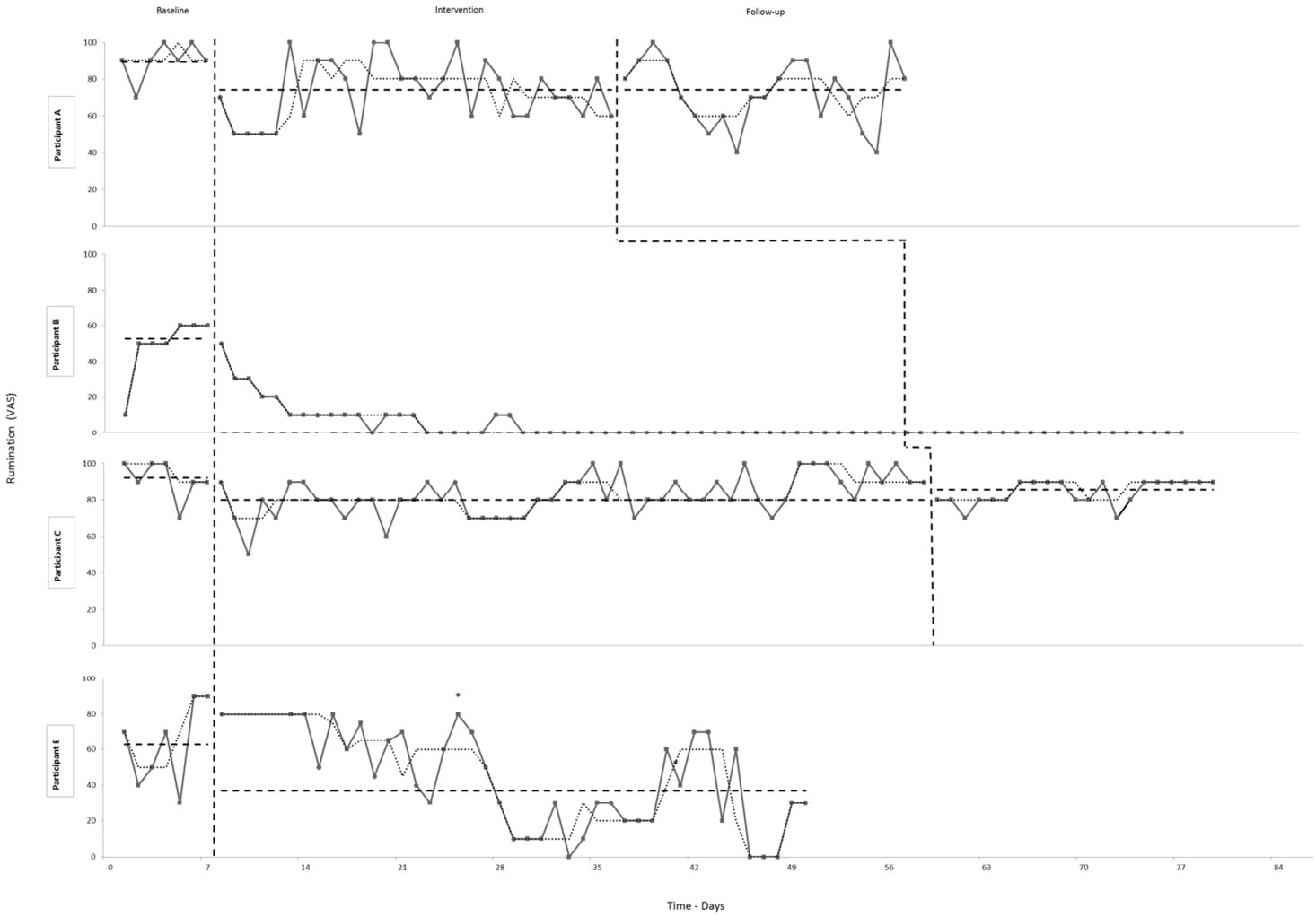



*Figure 4.* Rumination VAS data from participants A, B, C and E.

*Note.* Raw data = ; central tendency (broadened median) = ; trend (running median) = ; = session; * = life stressor. Participant E lost a friend to suicide and reported that it made him feel less alone abnormal. Different calculations of central tendency and trend plots were chosen according to Morley's visual analysis guidelines (Morley, 2015b).

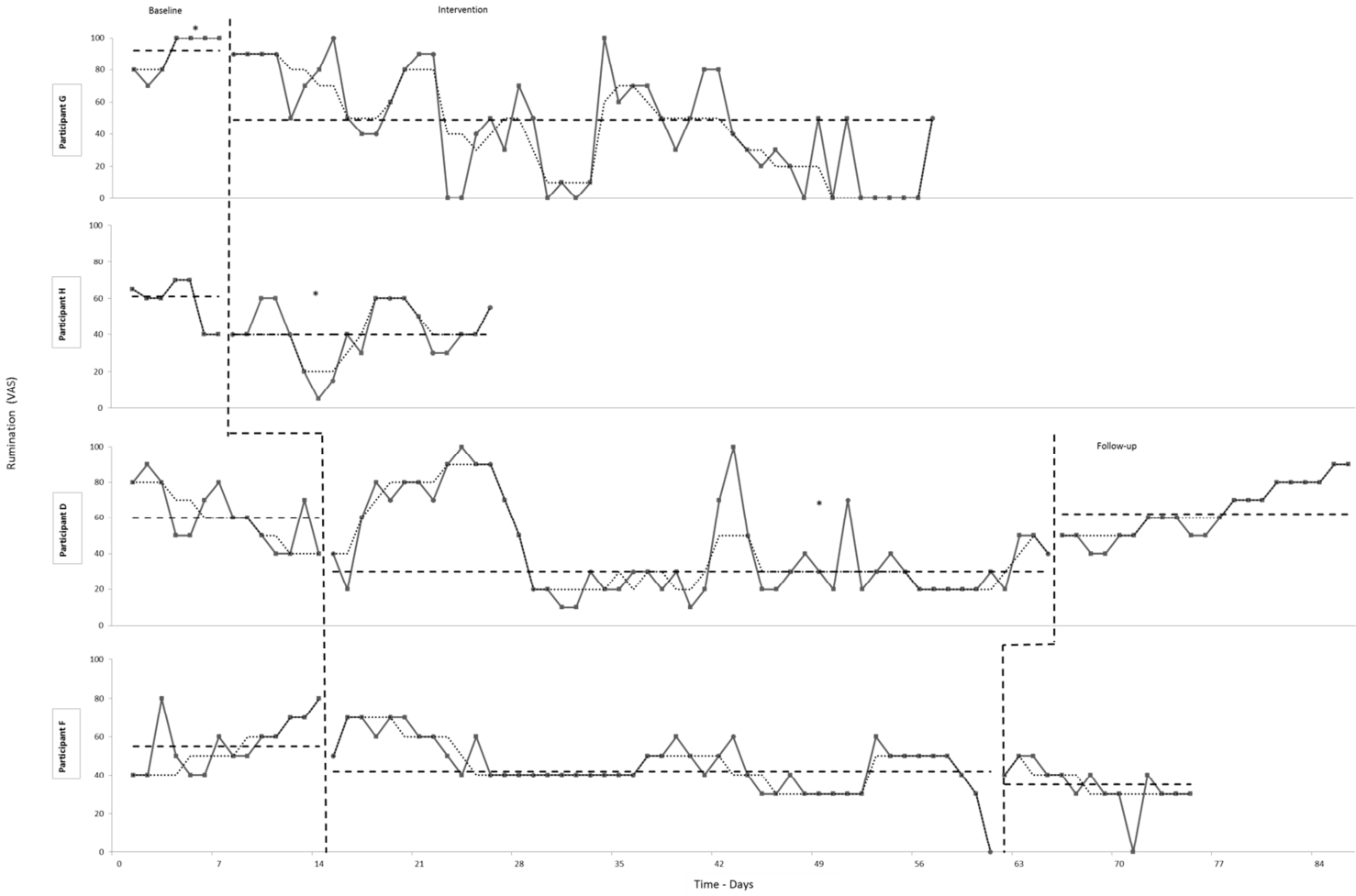



*Figure 5.* Rumination VAS data from participants G, H, D and F.

*Note.* Raw data = ■; central tendency (broadened median) = - - - - ; trend (running median) = ····· ; ●= session; * = life stressor. Participant D experienced a trauma, Participant G experienced a relationship breakdown and participant H stopped taking his medication. Different calculations of central tendency and trend plots were chosen according to Morley's visual analysis guidelines (Morley, 2015b).

Table 3 *Summary of tau analyses comparing idiographic depression ratings across the study phases.*

| Participant | Comparison | *Tau* | *SD Tau* | *p* value | 90% *CI* |
|---|---|---|---|---|---|
| A | A x B | -0.51 | 0.25 | .04* | [-0.91, -0.10] |
| | B x C | -0.18 | 0.17 | .28 | [-0.46, 0.09] |
| | A x (B+C) | -0.55 | 0.24 | .02* | [-0.94, -0.16] |
| B | A x B | -0.94 | 0.24 | <.001*** | [-1.00, -0.55] |
| | B x C | 0.25 | 0.15 | .10 | [0.00, 0.50] |
| | A x (B+C) | -0.96 | 0.23 | <.001*** | [-1.00, -0.58] |
| C | A x B | -0.31 | 0.23 | .19 | [-0.70, 0.08] |
| | B x C | -0.14 | 0.15 | .36 | [-0.39, 0.11] |
| | A x (B+C) | -0.36 | 0.23 | .11 | [-0.74, 0.01] |
| D | A x B | -0.55 | 0.18 | <.01** | [-0.84, -0.26] |
| | B x C | 0.80 | 0.15 | <.001*** | [0.55, 1.00] |
| | A x (B+C) | -0.21 | 0.17 | .21 | [-0.49, 0.07] |
| E | A x B | -0.66 | 0.24 | <.01** | [-1.00, -0.26] |
| F | A x B | -0.60 | 0.18 | <.001*** | [-0.89, -0.31] |
| | B x C | 0.34 | 0.18 | .05 | [0.05, 0.63] |
| | A x (B+C) | -0.65 | 0.17 | <.001*** | [-0.94, -0.37] |
| G | A x B | -0.83 | 0.24 | <.001*** | [-1.00, -0.45] |
| H | A x B | -0.80 | 0.26 | <.01** | [-1.00, -0.35] |
| Weighted average | A x B | -0.64 | | <.001*** | [-0.80, 0.00] |
| | B x C | 0.22 | | <.01** | [0.08, 0.36] |
| | A x (B+C) | -0.53 | | <.001*** | [-0.72, -0.35] |

*Note.* A = Baseline phase; B = Intervention phase; C = Follow-up phase; *CI = confidence interval*; *SD = standard deviation*; *Weighted averages* = combined effect-size; * = $p$ <.05; ** = $p$ <.01; *** = $p$ <.001

Table 4 *Summary of tau analyses comparing idiographic rumination ratings across the study phases.*

| Participant | Comparison | *Tau* | *SD Tau* | *p* value | 90% *CI* |
|---|---|---|---|---|---|
| A | A x B | -0.59 | 0.25 | .02* | [-0.99, -0.18] |
| | B x C | -0.00 | 0.17 | .98 | [-0.28, 0.27] |
| | A x (B+C) | -0.59 | 0.24 | .02* | [-0.98, -0.20] |
| B | A x B | -0.97 | 0.24 | <.001*** | [-1.00, -0.59] |
| | B x C | 0.23 | 0.15 | .13 | [-0.02, 0.487] |
| | A x (B+C) | -0.98 | 0.23 | <.001*** | [-1.00, -0.60] |
| C | A x B | -0.45 | 0.23 | .06 | [-0.83, -0.06] |
| | B x C | 0.10 | 0.15 | .52 | [-0.15, 0.34] |
| | A x (B+C) | -0.46 | 0.23 | .04* | [-0.84, -0.09] |
| D | A x B | -0.43 | 0.18 | .01* | [-0.72, -0.15] |
| | B x C | 0.53 | 0.15 | <.001*** | [0.28, 0.77] |
| | A x (B+C) | -0.29 | 0.17 | .09 | [-0.57, -0.01] |
| E | A x B | -0.45 | 0.24 | .06 | [-0.85, -0.06] |
| F | A x B | -0.48 | 0.18 | <.01** | [-0.77, -0.19] |
| | B x C | -0.46 | 0.18 | <.01** | [-0.76, -0.17] |
| | A x (B+C) | -0.55 | 0.17 | <.01** | [-0.83, -0.27] |
| G | A x B | -0.78 | 0.24 | <.001*** | [-1.00, -0.40] |
| H | A x B | -0.62 | 0.26 | .02* | [-1.00, -0.20] |
| Weighted average | A x B | -0.59 | | <.001*** | [-0.74, -0.40] |
| | B x C | 0.10 | | .18 | [-0.05, 0.23] |
| | A x (B+C) | -0.55 | | <.001*** | [-0.74, -0.38] |

*Note.* A = Baseline phase; B = Intervention phase; C = Follow-up phase; *CI = confidence interval*; *SD = standard deviation*; *Weighted averages* = combined effect-size; * = $p$ <.05; ** = $p$ <.01; *** = $p$ <.001

3.3. Reliable and clinically significant change

At the end of the intervention phase, seven participants had made reliable change on the MADRS (Table 5). Of these, three participants also met criteria for clinically significant change (B, C, & E). At follow-up, participant F continued to improve, whereas participant C experienced deterioration in her symptoms. Of those who completed follow-up data ($n$ = 5), three participants met criteria for reliable change (A, B, & F), of whom two also met criteria for clinically significant change (B & F). At the end of the intervention phase, four participants' scores demonstrated reliable change on the PHQ-9, three of whom also met criteria for clinically significant change (PB, PE, & PG). At follow-up, participants A and F showed continued improvement and overall three participants met criteria for reliable and clinically significant change (A, B, & F). Of note, some participants' responses clearly differed between the MADRS and PHQ-9. By the end of the intervention phase, all participants apart from participant C demonstrated reliable change in their BADS-SF scores. However, only participants B and E met criteria for clinically significant change on the BADS-SF. Following the assumption that overall treatment responders were those who demonstrated reliable and clinically significant change on one or both standardised measures of depression, half of the participants (B, C, E & G) were considered treatment responders at the end of their treatment, and three (60%) were considered treatment responders at follow-up (A, B & F).

Table 5 *Summary of reliable and clinically significant change in standardised outcomes*

| Participant | MADRS Baseline | MADRS End of treatment | MADRS Follow-up | PHQ-9 Baseline | PHQ-9 End of treatment | PHQ-9 Follow-up | BADS-SF Baseline | BADS-SF End of treatment |
|---|---|---|---|---|---|---|---|---|
| A | 39 | 23[RC] | 25[RC] | 15 | 13 | 10[CSC] | 7 | 28[RC] |
| B | 31 | 2[CSC] | 2[CSC] | 10 | 0[CSC] | 0[CSC] | 19 | 42[CSC] |
| C | 39 | 9[CSC] | 41 | 19 | 17 | 17 | 6 | 5 |
| D | 36 | 36 | 37 | 16 | 17 | 16 | 21 | 32[RC] |
| E | 30 | 5[CSC] | *n/a* | 19 | 7[CSC] | *n/a* | 12 | 43[CSC] |
| F | 53 | 19[RC] | 14[CSC] | 14 | 9 | 6[CSC] | 20 | 30[RC] |
| G | 45 | 18[RC] | *n/a* | 25 | 6[CSC] | *n/a* | 4 | 15 |
| H | 37 | 26 [RC] | *n/a* | 17 | 11[RC] | *n/a* | 18 | 26 |

*Note.* BAD-SF = Behavioural Activation Scale - Short Form (range: 0-54); MADRS = Montgomery Åsberg Depression Rating Scale (range: 0-60); PHQ-9 = Patient Health Questionnaire (range: 0-27); Reliable change ([RC]) and clinically significant change ([CSC]) indicated by Jacobson & Traux, 1991.

### 3.4. Client satisfaction and acceptability of the treatment

According to the CSQ, participants reported high treatment satisfaction ($M$ = 27.86, $SD$ = 4.49). Treatment acceptability ratings were also above average for both participants ($M$ = 81.43, $SD$ = 21.16) and therapists ($M$ = 66.67, $SD$ = 16.33), though noticeably higher for participants.

## 4. Discussion

Overall, this multiple baseline SCED has demonstrated that, time-intensive BA was feasible, and that, for the majority of participants ($n$ = 5 to 7) time-intensive BA was associated with significant improvements in idiographic symptoms of depression. Seven participants made reliable change on at least one standardised or process measure of depression, and at the end of their treatment, four participants were considered treatment responders. Most participants ($n$ = 3 or 4) who completed follow-up measures demonstrated either maintenance or improvement of progress on both idiographic and standardised measures of depression. Finally, treatment satisfaction and acceptability were rated as above average. As improvement on idiographic, standardised, and process outcome measures was shown for *some*, but not *all* participants, the research hypothesis is only partially accepted.

The findings add to the existing body of literature that supports the effectiveness of one session or time-intensive behavioural activation (Folke et al., 2015; Gawrysiak et al., 2009; Hopko et al., 2003a; Nasrin et al., 2017), but extends previous findings to BA (Martell et al., 2001; as opposed to BATD), delivered via multiple therapists, and to participants who met criteria for MDD. Promising findings might be explained by behavioural models of depression (Ferster, 1973; Lewinsohn,

1974) and BA's mechanism of extinguishing unhelpful behaviours that maintain depression, while increasing engagement in pleasant activities and response-contingent positive environmental reinforcement (Martell et al., 2001). However, specific mechanisms of change were not analysed here and require further investigation.

Importantly, the study's response rates are comparable to levels of responders obtained from the most recent non-inferiority trial of BA versus CBT (64%; Richards et al., 2016), and general depression treatment response rates found within IAPT services (55%; Richards & Borglin, 2011). These similarities could potentially be explained by the studies targeting similar UK populations and settings, or the limited sample size in the current study. However, the findings raise questions, which require further research, over whether time-intensive BA could be as effective as existing, empirically supported and recommended pacing of BA. As this study delivered less than the recommended amount of BA (NICE, 2009), yet still yielded mostly positive results, the findings support previous evidence to suggest that increasing treatment intensity promotes faster overall recovery (Erekson et al., 2015). It is foreseeable therefore, though not proven, that delivering time-intensive treatment could have scope for reducing service waiting list times.

It is interesting that level of treatment response, as well as which participants were considered responders, differed on the MADRS and the PHQ-9 at follow-up. Though moderators of effects were not examined here, discrepancies between these clinician and self-report rated measures could be explained by reporting bias, clinicians and participants having different standards for outcomes, or variation in the content and weighting of items of the MADRS and PHQ-9 (Cuijpers, Li, Hofmann &

Andersson, 2010), not to mention measurement error. Still, the discrepancies support evidence that clinician and self-report rated measures are not equivalent (Uher et al., 2012).

It was surprising that only two participants met criteria for *clinically significant* change in activation levels on the BADS-SF. However, these findings were consistent with those of Folke et al., (2015) who found that, despite their entire sample meeting reliable change on the BADS-SF, none made clinically significant change. Though only hypothetical at this point, it may be that our time-intensive BA does not allow enough between-session time for some participants to practice therapy skills to a sufficient level for clinically significant change.

Considering that attrition from existing depression treatments and BA studies has been higher than found here (Cooper & Conklin, 2015; Richards et al., 2016), the current finding of high treatment retention is encouraging, and tentatively suggests that delivering this BA format could discourage drop-out. However, the number of participants and therapists that were considered eligible or willing to participate in the study was limited by the inclusion/exclusion criteria. Some of the similarities of the participants' demographic characteristics (e.g., all being out of work and having had multiple previous episodes of depression) may also be a result of the inclusion/exclusion criteria, but these individuals could represent those most able to adhere to time-intensive BA. For example, time-intensive BA may be less feasible for working participants. In fact, a large proportion of participants approached to take part in the study did not have time to attend treatment time-intensively. This indicates that some persons with depression might prioritise not needing to take much time out of their existing routines to receive treatment and may also be indicative of the stigma

that continues to surround mental health.

Given that seven out of eight participants had received previous psychological therapy, it may be that time-intensive BA was attractive as a *new* and potentially *faster* treatment for recurrent depression, rather than simply as being more *accessible*. Clinicians intending to evaluate time-intensive BA should assess clients' motivations and remain vigilant to identify those potentially seeking a 'quick fix'.

The finding that some referrals declined involvement in the study due to a preference to receive CBT, suggests that the lay population may be less aware of BA and evidence suggesting its advantages in comparison to other treatments (Dimidjian et al., 2006; Dobson et al., 2008; Ekers et al., 2008; 2014; Richards et al., 2016; Sturmey, 2009), which this paper promotes, and could be promoted in future. It is understandable that treatments, when less well known and novel, might initially be adopted at a lower frequency.

Finally, low therapist recruitment rates, and therapists rating the acceptability of the intervention as lower than the participants did, suggests that IAPT may not be the most feasible setting for time-intensive treatment or efficacy research to take place. Components of the treatment design, such as daily prompting, may become overwhelming for therapists within certain settings. Services should carefully consider whether or not they can support time-intensive interventions or research into their efficacy. We encourage services to consider creative ways of enhancing how 'user-friendly' the treatment is , for example, by setting up text or email prompts to be sent automatically, protecting therapists session preparation time, and making time-intensive therapists exempt from other responsibilities of equal weighting. Undoubtably, person-centred approaches will need to be adopted to consider the

needs of the therapist and client, and discussion of costs and funding within the resource confines of each service would be necessary before trying out this new treatment modality. Therefore, the feasibility of time-intensive BA requires further research. Still, if time-intensive BA was adopted with high efficiency, and clients were treated faster (Erekson et al., 2015), one might envisage that in the long-run it would produce a corresponding reduction in workload for participating therapists.

The study findings are limited by only five participants completing a short follow-up phase, restricting conclusions that can be drawn about the treatment's durability. In addition, the generalizability of the findings is limited by the study's strict inclusion criteria, small, homogenous sample size, and only recruiting participants and therapists from three London IAPT services. Specifically, this meant that the sample was mainly made up of white females who were out of work, and the therapists were working within less flexible service structures. The reliability of collected data was subject to self-report biases, and experimenter bias where clinician-rated measures were not inter-rated for reliability. Furthermore, treatment fidelity cannot be assumed, as only five session tapes were assessed. Lastly, the non-concurrent treatment design is subject to the effects of confounding variables (e.g., maturation), reducing the study's internal validity. Indeed, participants' differing treatment sequences and reported stressors might have predicted some variation in outcome measures.

Future research should aim to rectify the limitations of the current study, primarily by replicating it across different samples and settings, and extending the follow-up phase. Considering that patterns of change on idiographic, standardised and process measures, of the same concept, were not always congruent, future research

should continue to collect all three outcome modalities. Within larger studies regression analyses should be conducted to determine whether or not treatment characteristics (e.g., duration, session number, session spacing, and prompting), service characteristics (e.g., IAPT or non-IAPT), participant characteristics (e.g., history, comorbidity and life-stressors), and therapist characteristics (e.g., experience and number) predict treatment responses. Coding frameworks should also be used to determine and track specific (e.g., prompting levels) and non-specific (e.g., therapeutic relationship strength) mechanisms of change that occur in each session, their temporal relationships to outcomes, and whether or not they are crucial to recovery following time-intensive BA. Qualitative research investigating reasons behind acceptability ratings should also be conducted. Such findings could then be used to guide the development of an optimized time-intensive BA. In the longer-term future, RCTs should be conducted to determine the efficacy and cost-effectiveness of different time-intensive BA designs in comparison to each other (e.g., 12 hours of BA delivered over one week in comparison to 12 hours of BA and prompting delivered across two weeks), control and/or recommended active control conditions (e.g., weekly BA or time-intensive CBT for depression). We anticipate that effectively treating depression via time-intensive treatment might still require longer treatment durations than some anxiety disorders have done; however, addressing these hypotheses to optimize time-intensive treatment would be worthwhile. Importantly, future research should adopt a shared definition of ‘time-intensive’, and treatment ‘responder’ in a bid to move towards standardisation of terminology and outcomes.

**Conclusions**

In summary, the current study provides new and tentative evidence highlighting the potential of time-intensive BA as a feasible, acceptable, and effective intervention for some, though not all, adult outpatients with depression. This constitutes an essential step in attempting to increase patient choice and access to depression treatments. The current findings now warrant further exploration in order to be substantiated. Once such further research has been completed, we will be more able to determine whether or not this treatment approach can be effectively disseminated within outpatient settings to promote the well-being of the population.

**References**

Bloom, D., Cafiero, E., Jané-Llopis, E., Abrahams-Gessel, S., Bloom, L., Fathima, S., ... & O'Farrell, D. (2012). *The global economic burden of noncommunicable diseases*. Geneva: World Economic Forum.

Cooper, A. A., & Conklin, L. R. (2015). Dropout from individual psychotherapy for major depression: A meta-analysis of randomized clinical trials. *Clinical Psychology Review, 40*, 57-65.

Cuijpers, P. (2015). Psychotherapies for adult depression: Recent developments. *Current Opinion in Psychiatry, 28*(1), 24-29.

Cuijpers, P., Huibers, M., Ebert, D. D., Koole, S. L., & Andersson, G. (2013). How much psychotherapy is needed to treat depression? A meta-regression analysis. *Journal of Affective Disorders, 149*(1), 1-13.

Cuijpers, P., Li, J., Hofmann, S. G., & Andersson, G. (2010). Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy

research on depression: A meta-analysis. Clinical Psychology Review, 30(6), 768-778.

Cunningham, J. L., Wernroth, L., von Knorring, L., Berglund, L., & Ekselius, L. (2011). Agreement between physicians' and patients' ratings on the Montgomery–Åsberg Depression Rating Scale. *Journal of Affective Disorders, 135*(1), 148-153.

Department of Health (DOH). (2008). *Improving Access to Psychological Therapies: Implementation plan: National guidelines for regional delivery.* Retrieved from http://socialwelfare.bl.uk/subject-areas/services-client-groups/adults-mental-health/departmentofhealth/improving08.aspx

Dimidjian, S., Barrera, J. M., Martell, C., Muñoz, R. F., & Lewinsohn, P. M. (2011). The origins and current status of behavioral activation treatments for depression. *Annual Review of Clinical Psychology, 7*, 1-38.

Dimidjian, S., Goodman, S. H., Sherwood, N. E., Simon, G. E., Ludman, E., Gallop, R., . . . Hubley, S. (2017). A pragmatic randomized clinical trial of behavioral activation for depressed pregnant women. *Journal of Consulting and Clinical Psychology, 85*(1), 26-36.

Dimidjian, S., Hollon, S. D., Dobson, K. S., Schmaling, K. B., Kohlenberg, R. J., Addis, M. E., . . . Gollan, J. K. (2006). Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *Journal of Consulting and Clinical Psychology, 74*(4), 658-670.

Dimidjian, S., Hubley, S., Martell, C. R., & Herman, R. (2016). *Quality of Behavioral Activation Scale* (Unpublished manuscript). University of Colorado Boulder, USA.

Dimidjian, S., Martell, C. R., Addis, M. E., & Herman-Dunn, R. (2008). Behavioral activation for depression. In D. H. Barlow (Ed.) *Clinical handbook of psychological disorders: A step-by-step treatment manual (*pp.328-364). New York: Guildford Press.

Dobson, K. S., Hollon, S. D., Dimidjian, S., Schmaling, K. B., Kohlenberg, R. J., Gallop, R. J., . . . Jacobson, N. S. (2008). Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the prevention of relapse and recurrence in major depression. *Journal of Consulting and Clinical Psychology, 76*(3), 468-477.

Ehlers, A., Hackmann, A., Grey, N., Wild, J., Liness, S., Albert, I., . . . Clark, D. M. (2014). A randomized controlled trial of 7-day intensive and standard weekly cognitive therapy for PTSD and emotion-focused supportive therapy. *American Journal of Psychiatry, 171*(3), 294-304.

Erekson, D. M., Lambert, M. J., & Eggett, D. L. (2015). The relationship between session frequency and psychotherapy outcome in a naturalistic setting. *Journal of Consulting and Clinical Psychology, 83*(6), 1097-1107.

Ferrari, A. J., Charlson, F. J., Norman, R. E., Patten, S. B., Freedman, G., Murray, C. J., . . . Whiteford, H. A. (2013). Burden of depressive disorders by country, sex, age, and year: Findings from the global burden of disease study 2010. *PLoS Medicine, 10* (11). Retrieved from https://doi.org/10.1371/journal.pmed.1001547

Ferster, C. B. (1973). A functional analysis of depression. *American Psychologist, 28*(10), 857-870.

First, M. B., Williams, J. B. W., Karg, R. S., & Spitzer, R. L. A. (2015). *Structured Clinical Interview for DSM-5 Disorders, Clinician Version (SCID-5-CV).* Virginia, VA: American Psychiatric Association.

Folke, F., Hursti, T., Tungström, S., Söderberg, P., Kanter, J. W., Kuutmann, K., . . . Ekselius, L. (2015). Behavioral activation in acute inpatient psychiatry: A multiple baseline evaluation. *Journal of Behavior Therapy and Experimental Psychiatry, 46*, 170-181.

Gawrysiak, M., Nicholas, C., & Hopko, D. R. (2009). Behavioral activation for moderately depressed university students: Randomized controlled trial. *Journal of Counseling Psychology, 56*(3), 468-475.

Glover, G., Webb, M., & Evison, F. (2010). *Improving access to psychological therapies: A review of the progress made by sites in the first rollout year*. UK: North East Public Health Observatory.

Herrnstein, R. J. (1970). On the law of effect. *Journal of The Experimental Analysis of Behavior, 13*(2), 243-266.

Hopko, D. R., Lejuez, C., Lepage, J. P., Hopko, S. D., & McNeil, D. W. (2003a). A brief behavioral activation treatment for depression: A randomized pilot trial within an inpatient psychiatric hospital. *Behavior Modification, 27*(4), 458-469.

Hopko, D. R., Lejuez, C., Ruggiero, K. J., & Eifert, G. H. (2003b). Contemporary behavioral activation treatments for depression: Procedures, principles, and progress. *Clinical Psychology Review, 23*(5), 699-717.

Jacobson, N., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. . *Journal of Consulting and Clinical Psychology, 59*(1), 12-19.

Kanter, J. W., Manos, R. C., Bowe, W. M., Baruch, D. E., Busch, A. M., & Rusch, L. C. (2010). What is behavioral activation?: A review of the empirical literature. *Clinical Psychology review, 30*(6), 608-620.

Kanter, J. W., Puspitasari, A. J., Santos, M. M., & Nagy, G. A. (2012). Behavioural activation: History, evidence and promise. *The British Journal of Psychiatry, 200*(5), 361-363.

Kazdin, A. E. (1982). *Single-case experimental designs: Strategies for studying behavior change*. New York: Oxford University Press.

Kazdin, A. E. (1998). *Research design in clinical psychology.* Needham Heights, MA: Allyn & Bacon.

Kratochwill, T., Hitchcock, J., Horner, R., Levin, J. R., Odom, S., Rindskopf, D., & Shadish, W. (2010). *Single-case designs technical documentation*. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: A new depression diagnostic and severity measure. *Psychiatric Annals, 32*(9), 509-515.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The Phq□9. *Journal of General Internal Medicine, 16*(9), 606-613.

Larsen, D. L., Attkisson, C. C., Hargreaves, W. A., & Nguyen, T. D. (1979). Assessment of client/patient satisfaction: Development of a general scale. *Evaluation and Program Planning, 2*(3), 197-207.

Lejuez, C., Hopko, D. R., LePage, J. P., Hopko, S. D., & McNeil, D. W. (2001). A brief behavioral activation treatment for depression. *Cognitive and Behavioral Practice, 8*(2), 164-175.

Lewinsohn, P. M. (1974). Clinical and theoretical aspects of depression. In K. S. Calhoun (Ed.), *Innovative Treatment Methods in Psychopathology* (Vol. 1, pp. 63-120). New York: Wiley.

Manos, R. C., Kanter, J. W., & Luo, W. (2011). The behavioral activation for depression scale–short form: Development and validation. *Behavior Therapy, 42*(4), 726-739.

Martell, C. R., Addis, M. E., & Jacobson, N. S. (2001). *Depression in context: Strategies for guided action*. New York: WW Norton & Co.

Martell, C. R., Dimidjian, S., & Herman-Dunn, R. (2013). *Behavioral activation for depression: A clinician's guide*. New York: Guilford Press.

McCrone, P., Dhanasiri, S., Patel, A., Knapp, M., & Lawton-Smith, S. (2008). *Paying the price: The cost of mental health care in England to 2026.* London: King's Fund.

McMillan, D., Gilbody, S., & Richards, D. (2010). Defining successful treatment outcome in depression using the PHQ-9: A comparison of methods. *Journal of Affective Disorders, 127*(1), 122-129.

Mohr, D. C., Ho, J., Duffecy, J., Baron, K. G., Lehman, K. A., Jin, L., & Reifler, D. (2010). Perceived barriers to psychological treatments and their relationship to depression. *Journal of Clinical Psychology, 66*(4), 394-409.

Montgomery, S. A., & Åsberg, M. (1979). A new depression scale designed to be sensitive to change. *The British Journal of Psychiatry, 134*(4), 382-389.

Moran, P., Leese, M., Lee, T., Walters, P., Thornicroft, G., & Mann, A. (2003). Standardised Assessment of Personality-Abbreviated Scale (SAPAS): Preliminary validation of a brief screen for personality disorder. *The British Journal of Psychiatry, 183*(3), 228-232.

Morley, S. (2015a). Single cases are complex. *Scandinavian Journal of Pain, 7*, 55-57.

Morley, S. (2015b). Visual analysis for single-case data: Draft chapter. Retrieved from https://www.researchgate.net/publication/278300494_Visual_analysis_for_single_case_data_Draft_chapter.

Nasrin, F., Rimes, K., Reinecke, A., Rinck, M., & Barnhofer, T. (2017). Effects of brief behavioural activation on approach and avoidance tendencies in acute depression: Preliminary findings. *Behavioural and Cognitive Psychotherapy, 45*(1), 58-72.

National Institute for Health and Clinical Excellence (NICE). (2009). *Depression in adults: Recognition and management*. UK: National Institute for Health and Clinical Excellence.

Parker, R. I., & Vannest, K. J. (2012). Bottom-up analysis of single-case research designs. *Journal of Behavioral Education, 21*(3), 254-265.

Parker, R. I., Vannest, K. J., Davis, J., & Sauber, S. (2011). Combining non-overlap and trend for single-case research: Tau-U. . *Behavior Therapy, 42*, 284-299.

Read, A., Mazzucchelli, T. G., & Kane, R. T. (2016). A preliminary evaluation of a single session behavioural activation intervention to improve well□being and prevent depression in carers. Clinical Psychologist, 20(1), 36-45.

Richards, D. A., & Borglin, G. (2011). Implementation of psychological therapies for anxiety and depression in routine practice: Two year prospective cohort study. *Journal of Affective Disorders, 133*(1), 51-60.

Richards, D. A., Ekers, D., McMillan, D., Taylor, R. S., Byford, S., Warren, F. C., . . . Kuyken, W. (2016). Cost and outcome of behavioural activation versus

cognitive behavioural therapy for depression (COBRA): A randomised, controlled, non-inferiority trial. *The Lancet*, *388*(10047), 871-880.

Santos, M. M., Rae, J. R., Nagy, G. A., Manbeck, K. E., Hurtado, G. D., West, P., . . . Kanter, J. W. (2017). A client-level session-by-session evaluation of behavioral activation's mechanism of action. Journal of Behavior Therapy and Experimental Psychiatry, 54, 93-100.

Storch, E. A., Geffken, G. R., Merlo, L. J., Mann, G., Duke, D., Munson, M., . . . Goodman, W. K. (2007). Family-based cognitive-behavioral therapy for pediatric obsessive-compulsive disorder: Comparison of intensive and weekly approaches. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*(4), 469-478.

Svanborg, P., & Åsberg, M. (1994). A new self□rating scale for depression and anxiety states based on the Comprehensive Psychopathological Rating Scale. *Acta Psychiatrica Scandinavica*, *89*(1), 21-28.

Uher, R., Perlis, R. H., Placentino, A., Dernovšek, M. Z., Henigsberg, N., Mors, O., . . . Farmer, A. (2012). Self□report and clinician□rated measures of depression severity: Can one replace the other? *Depression and Anxiety*, *29*(12), 1043-1049.

Zlomke, K., & Davis, T. E. (2008). One-session treatment of specific phobias: A detailed description and review of treatment efficacy. *Behavior Therapy*, *39*(3), 207-223.

Highlights

- Eight participants with MDD completed time-intensive behavioural activation.
- Six participants made significant improvements in their idiographic symptoms.
- Four participants were considered treatment responders.
- Follow-up data showed that treatment effects were mainly maintained.
- Participants and therapists considered the intervention to be acceptable.

**Acknowledgements, declaration of interest and role of funding organizations**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Salary support is provided to Professor David Veale by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. We would also like to thank the therapists who were involved in delivering the intervention: Isabel Antis, Adam Corfe, Lisanne Davies, Dr Claudia Hallett, Amy Holloway, Gareth Stephens, Jennifer Scott, and Dr Sharif Elithy, for his expert advice.