Bringing light into darkness: A multiple baseline mixed methods case series evaluation of Augmented Depression Therapy (ADepT)

Barnaby D Dunn[1*]

Emily Widnall[1]

Nigel Reed[1]

Christabel Owens[2]

John Campbell[2]

Willem Kuyken[3]

1 - Mood Disorders Centre, University of Exeter

2 - College of Medicine and Health, University of Exeter

3 - Department of Psychiatry, University of Oxford

*=corresponding author. Barney Dunn, Mood Disorders Centre, University of Exeter, EX4 4QG, b.d.dunn@exeter.ac.uk

Highlights

- A case series evaluated Augmented Depression Therapy (ADepT)
- ADepT was acceptable to treatment resistant depressed clients.
- ADepT was effective at reducing depression and building wellbeing.
- ADepT may be superior to existing treatments at reducing anhedonia.

Abstract

(194 words)

Two core features of depression are elevations in negative valence system (NVS) functioning and reductions in positive valence system (PVS) functioning. Existing psychological treatments have focused on the NVS and neglected the PVS, which may contribute to sub-optimal outcomes. The present mixed methods multiple randomised baseline case series preliminarily evaluates Augmented Depression Therapy (ADepT), a novel depression treatment targeting PVS and NVS disturbance that aims both to reduce depression and enhance wellbeing. Eleven clinically depressed participants were recruited. Intensive time series analyses showed that 7/11 participants improved on both wellbeing and depression. Reliable and clinically significant improvement was observed for 9/11 participants on at least one of these outcomes (and also across a range of other PVS and NVS outcomes). Group level analyses showed significant pre to post change on all outcomes. Benchmarking analyses indicated these effect sizes were at least comparable (and for some PVS outcomes superior) to existing treatments. Gains were largely sustained over one-year follow-up. Qualitative interviews indicated ADepT was feasible and acceptable. These findings provide preliminary support for ADepT as a novel depression treatment. Further evaluation directly comparing ADepT to existing treatments using randomised controlled trial designs is now required.

Depression is a prevalent, recurrent and debilitating condition, leads to marked economic and social costs, and is predicted to become the leading worldwide contributor to disability by 2020. (Judd, 1997; Singleton et al., 2003; Kruijshaar et al., 2005; Üstün et al., 2004; Andrews & Titov, 2007; Moussavi et al., 2007; Kessler et al., 2003; Layard et al., 2006). Existing psychological and pharmacological interventions for depression are not optimally effective, with approximately only half of individuals responding (showing at least a 50% reduction in symptoms) and many exhibiting residual depression symptoms and functional impairment after treatment (Cuijpers et al., 2008; Rush, 2015; Sheehan et al., 2011). It remains unclear if existing treatments are superior to placebo, except in cases of severe depression (Cuijpers et al., 2014a, 2014b; Fournier et al., 2010). Even in those who fully recover during acute treatment, subsequent rates of relapse are high (Vittengel et al., 2007). Enhanced treatments are needed to reduce symptom severity and to improve functionality, allowing individuals to lead the best life they can alongside depression.

One way forward is to acknowledge that depression may not be a homogeneous diagnostic entity and instead may be better considered in terms of a constellation of distinct underlying functional dimensions, each requiring different intervention strategies (Insel et al., 2010). The two symptoms required for a diagnosis of depression are either a pervasive depressive mood or a loss of pleasure and interest in all or most activities (anhedonia). These symptoms can be seen as emergent properties of disruptions to two underlying and partly dissociable neurobiological dimensions in depression (Watson, Wiese, Vaidya, & Tellegen, 1999; Gray, 1987; Carver & White, 1994; Paulus et al., 2017; Insel et al., 2010). Up-regulation of the negative valence system (NVS) results in elevated withdrawal from punishing stimuli and increased negative affect (NA) experience. Down-regulation of the positive valence system (PVS) reduces approach to rewarding stimuli and inhibits positive affect (PA) experience.

Whilst it has long been recognised that NVS elevations are prognostically important in depression, it is now increasingly acknowledged that PVS reductions also predict acute symptom severity, poor treatment response, and a chronic, relapsing future course (Spijker, Bijl, De Graaf, & Nolen, 2001; Uher et al., 2012, McMakin et al., 2012; see reviews by Dunn & Roberts, 2016; Dunn, in press). To treat depression effectively it is therefore likely to be necessary to target disturbances in both the PVS and the NVS. However, existing psychological treatments primarily focus on NVS elevations and neglect the PVS (Dunn,

2012; Dunn & Roberts, 2016). For example, informal content analysis suggests that Cognitive Therapy (CT; Beck et al., 1979) predominately targets the modification of negative thinking to reduce negative mood (Dunn, in press).

Consistent with the assertion that existing approaches have neglected the PVS, treatment has been shown to be more effective at lowering NA than building PA in routine practice (Brown, 2007; Naragon-Gainey, Gallagher & Brown, 2013; Kring, Persons & Thomas, 2007). Similar findings emerge when reanalysing existing RCTs (Dunn et al., submitted). The CPT2 trial (DeRubeis et al., 2005) compared 16 weeks of cognitive therapy and antidepressant medication in the treatment of depression and the CPT3 trial (Hollon et al., 2014) investigated the efficacy of antidepressant medication alone versus combined cognitive therapy and antidepressant medication in treating depression to remission. Secondary analyses revealed that reductions in PA were more marked than elevations in NA prior to treatment; there was a smaller repair of PA relative to NA during treatment; and disturbances were greater in PA versus NA at the end of treatment (Dunn et al., submitted). A secondary analysis of the COBRA trial (Richards et al., 2016) has also examined how well CT and Behavioural Activation (BA) repair anhedonia symptoms (Dunn et al., in prep b). Neither CT nor BA treatment fully repaired anhedonia, meaning at the end of treatment participants still showed elevated levels of anhedonia[1].

To improve depression outcomes, enhanced treatments need to be developed that simultaneously target both the PVS and the NVS (Dunn, 2012; Dunn & Roberts, 2016; Wood & Tarrier, 2010), focusing on underlying psychological mechanisms that exacerbate NA and reduce PA. When deciding how to measure treatment outcomes, it is also imperative to consider what 'recovery' from depression means from a client perspective. The key element of recovery from a client perspective is enhanced functioning in valued life domains (i.e. a reduction in the disability burden of depression) and increased wellbeing (the capacity to experience pleasure, meaning and social connection in life; Keyes 2005) (Zimmerman et al., 2006; Demyttenaere et al., 2015). This resonates with the 'recovery' literature emphasising that recovery involves a positive focus on being able to live a life that enhances wellbeing and functioning (Slade, 2010), often alongside mental illness (Anthony, 2003). Given the chronic, relapsing course of depression, focusing on functional recovery of this kind rather than complete cure may feel a more realistic goal for many clients.

The goal-setting literature also emphasises the importance of identifying approach rather than avoidance goals (Elliot, Sheldon & Church, 1997; Roskes, Elliot & DeCru, 2014), meaning it may be preferable to target building wellbeing rather than symptom relief. It is known that functional improvement can often lag behind symptomatic improvement (Rush, 2015; Sheehan et al., 2011; Sheehan et al., 2017). Therefore, enhanced treatments need to target both symptom relief *and* enhanced functioning and wellbeing, helping individuals learn to live well alongside depression. Such an approach is likely to have high acceptability to clients, potentially reducing drop-out during treatment (estimated to be greater than 25% during CT; Hans & Willer, 2013; Fernandez, Salem, Swift & Ramtahal, 2015).

Augmented Depression Therapy (ADepT) has been developed to target PVS and NVS deficits in depression simultaneously. ADepT is a fifteen session individual therapy for acute depression (with five optional booster sessions in the year post treatment). To help foster positive recovery, ADepT co-targets reduction of depression symptoms *and* enhancement of wellbeing/functioning. Following MRC guidance regarding the development of complex interventions (Craig et al., 2008), the ADepT protocol was constructed by: i) co-designing the intervention with service-users and other key stakeholders to maximise uptake (Muller, 2012; Cooper, 1999; Concannon et al., 2012); ii) translating findings from basic science research characterising PVS deficits in depression (cf., Clark, 2004); and iii) integrating elements of best practice from existing treatment approaches. The co-design process followed the principles of Intervention Mapping (Bartholomew et al., 2011) and involved an iterative process of stakeholder consultation, literature review and analysis of the local context, during which a preliminary logic model of the intervention was tested and revised. We used proven elements from existing treatments to target NVS function. Basic science findings were used to identify those underlying mechanisms impairing PVS function which ADepT should target (cf. Dunn, 2017). The mechanisms targeted included: elevated use of dampening appraisals when feeling positive (Burr, Javiad, Jell, Werner-Seidler & Dunn, 2017, Dunn et al., 2018; Werner-Seidler et al., 2013) ; reduced attentional and mnemonic processing biases to positive information (e.g., see Winer & Salem, 2016; Matt, Vazquez & Campbell, 1992; Dunn et al., 2009); reductions in experiential processing (e.g., see Gadeikis et al., 2017; Teasdale, 1999); and reductions in self-compassion and kindness (e.g. see

Hofmann, Grossman & Hinton, 2011). These mechanisms can all be seen as subtypes of avoidance of positive emotion experience (e.g. Jacob, Ower & Bucholz, 2013; Bardeen et al., 2014; Gilbert et al., 2014).

An effective and efficient first step in evaluating and refining a novel psychological treatment is to conduct a case series (Barlow, Nock & Hersen, 2008; Kazdin, 2011; Morley, 2017; Levin & Bressler, 1996; Salkovskis, 1995). Case series methods can assess treatment acceptability and feasibility; provide preliminary evidence of effectiveness and proof-of-concept; help refine and optimise the protocol; and establish if the treatment is likely to be implementable in the target context. By replicating findings across a series of individual cases, this begins to control for threats to external validity and to assess whether findings are generalizable. Intensive time series data are typically collected and analysed at the individual participant level to ascertain visually and statistically whether there is within-person change in slope and level from a baseline phase to the intervention. Where a sufficient number of data points are collected, these individual time series analyses have adequate power to test statistically the efficacy of an intervention for an individual participant (Borckardt et al., 2008). Use of a randomised multiple baseline design (randomising individuals to different lengths of baseline phase before starting treatment; effectively turning case series into rigorous experimental design) further controls to threats to validity. In particular, it helps to differentiate between a genuine treatment effect and natural recovery over time or other confounding factors (Kratochwill & Levin, 2010). These time series analyses can also be supplemented with reliably and clinically significant change analyses on standardized assessment measures. These can determine the proportion of individuals who improve to a greater extent than measurement error on the scale and beyond a minimum amount that indicates a meaningful degree of change (Jacobson & Truax, 1991). Case series can also highlight treatment non-responders, who may otherwise be hidden within the overall effect of a group-based design.

Given the lack of an internal control condition in case series designs, it can be helpful to benchmark the treatment effects against some external comparator (Borckardt et al., 2008). Conventional group level statistical tests can be run to estimate effect sizes, which can be compared against existing randomised control trial (RCT) outcomes for other depression interventions. Case series can be further enhanced by incorporating qualitative methods, allowing for detailed exploration of patient views on feasibility, acceptability, efficacy, mechanisms of action and ways in which the treatment can be improved (O'Cathain,

Thomas, Drabble, Rudolph & Hewison, 2013; Onghena, Maes & Heyvaert, 2018). While quantitative analyses can provide numerical precision, qualitative analyses can ensure descriptive precision (Kichenhaum, 2010; Onghena et al., 2018). Further, alongside a careful examination of participant characteristics and extra-therapeutic events, qualitative analysis can help identify reasons why a treatment did not produce change in a particular individual and for whom a treatment might be contraindicated.

We therefore ran a randomised multiple baseline case series evaluation of ADepT, including both quantitative and qualitative methods. We set *a priori* continuation rules allowing us to proceed to pilot trial evaluation of ADepT only if:

1. 60% of clients show at least reliable improvement pre to post in depression or wellbeing
2. < 30% of clients show reliable clinical deterioration pre to post in depression or wellbeing
3. > 60% of clients complete a minimum adequate dose of acute treatment (over 50% of acute dose; 8 or more sessions)
4. Clients on average attend > 60% of scheduled acute sessions (on average at least 7.5 sessions per client)
5. > 60% of clients and therapists rate treatment as acceptable, that they are satisfied with the treatment, and that they would recommend it to friends and family
6. ADepT treatment participation does not lead to serious negative consequences for participants (unexpected, clearly trial- or treatment-related serious adverse reaction).
7. At least large pre to post effect sizes emerge on depression and wellbeing (Hedges g > 0.8; cf. Cohen, 1988).

Further, the case series outcomes were benchmarked against existing trial data to examine how NVS (depression and anxiety symptoms, negative affect) and PVS (positive affect, anhedonia, and wellbeing) outcomes compare to existing treatments. While this was not pre-specified as a continuation rule, the expectation was that ADepT would lead to broadly comparable NVS effects, and superior PVS effects, relative to current best practice.

Method

Setting and Design

The case series was conducted at the Accessing Evidence Based Psychological Therapies (AccEPT) clinic, University of Exeter (see: https://www.exeter.ac.uk/mooddisorders/acceptclinic/), utilising a randomised multiple baseline mixed methods case-series design. Participants were recruited by writing to clients on local Improving Access to Psychological Therapy (IAPT) Services high intensity waiting lists for depression therapy or via NHS or self-referral to the AccEPT clinic. Participants were randomised to different baseline assessment length (between three and eight weeks), with a random sequence generated by a computer-based package. The acute intervention phase consisted of up to fifteen weekly sessions. Participants completed measures of depression and wellbeing each week in the baseline and acute treatment phase. In addition, participants also completed a longer battery of measures at randomisation, pre-treatment, post-treatment, and two months after completing acute treatment. An ethics amendment was sought after the original study had completed to follow-up participants approximately one year later (after the optional booster sessions had been completed). We followed guidelines in the design, analysis and reporting of case series to ensure the method was applied optimally (Smith, 2012; Kratochwill et al., 2010, 2012; Tate et al., 2016).

Participants

Consecutive referrals that met study inclusion criteria took part in the study. Inclusion criteria were being aged over 18 years; currently experiencing a major depressive episode based on a Structured Clinical Interview for Diagnosis (SCID-I; First et al.,1994); and scoring in the clinical range of the PHQ-9 with marked anhedonic features (PHQ-9 total score$\geq$ 10; item one measuring anhedonia score $\geq$2). Participants were required to describe depression as their primary presenting problem and have sufficient knowledge of written and spoken English to be able to make use of the therapy and to complete research assessments without the need of a translator. Exclusion criteria included: a history of bipolar disorder or organic brain change, currently receiving any other psychosocial therapy, substance abuse that compromised ability to use therapy, and current marked risk to self (self-harm or suicide) that could not be safely managed in the clinic

setting. Taking psychotropic medication was not an exclusion criterion in this study and a majority of participants were prescribed anti-depressants.

Sample size estimations for intensive time series analysis (our primary analysis method) do not follow a formal power analysis approach. However, guidelines recommend a minimum of three replications of the intervention effect (Kratochwill et al., 2010). To increase confidence in the generalisability of the findings, we asked three different therapists to deliver the treatment, aiming for each to treat at least three cases. For the secondary group level analyses (used to generate effect sizes for benchmarking purposes), we aimed to be powered to detect a large effect size in the key pre-post analysis (based on the continuation rule that ADepT needed to demonstrate a large effect size on the primary outcome variables). We therefore intended to recruit at least 13 participants. We only included participants in the final analyses who remained above the clinical cut-off for depression at the end of the baseline phase (PHQ-9$\geq$10) and showed a stable pattern of depression and wellbeing during the baseline phase. This is because instability in the baseline phase can bias the conclusions reached in both visual and statistical analyses (Kratochwill et al., 2010). We used reliable change for the PHQ-9 ($\geq$ 6 points) and wellbeing ($\geq$ 7 points) as the cut-off to indicate the baseline was not stable.

Participants gave written informed consent prior to participation. The original study (and the amendment to follow-up participants at one year) was approved by Frenchay NHS ethics committee (reference number 15/SW/0352). Participants were given an honorarium of £10 for each of the research assessments they completed (a maximum of £60).

Intervention

ADepT is an integrationist approach that combines novel intervention elements translated from basic science alongside existing strategies from established treatments, including CT (Beck, 1979; Moore & Garland, 2003; Kuyken, Padesky & Dudley, 2009), Acceptance and Commitment Therapy (ACT; Hayes, Strosahl & Wilson, 2011), Behavioural Activation (BA; Martell, Dimidjian & Herman-Dunn, 2013), Mindfulness Based Cognitive Therapy (MBCT; Segal, Williams & Teasdale, 2012), and positive psychology approaches (Sin & Lyubomirsky, 2009; Bryant & Veroff, 2007).

ADepT aims to enhance wellbeing (positive mood, meaning and social connection), to optimise functioning and to reduce symptoms of depression. It is a solution-focused, values-based, cognitively augmented behavioural activation approach, consisting of up to 15 acute treatment sessions and up to 5optional booster sessions (scheduled flexibly in the year following acute treatment to sustain longer-term recovery). Each session lasts 60 minutes, with the exception of the initial assessment which can last up to 90 minutes. Sessions are scheduled weekly and are delivered face-to-face, although therapists are encouraged to be flexible about timing and delivery format to meet client needs. Each session follows a typical CT structure consisting of: mood review, agenda setting, home practice review, working through each agenda item, and client summary and feedback. Figure One presents the logic model underpinning ADepT (see also Dunn et al., 2019 for further details of ADepT protocol).

Therapist Style

Therapists are encouraged to make judicious use of warmth, humour, positive feedback, gratitude and self-disclosure to build a positive relationship with clients and to make each session enjoyable where possible (logic model box 1). A solution-focused, coaching stance is taken, intentionally focusing on pockets of resilience and wellbeing to enhance client agency whenever possible ('thickening the positive narrative'; De Shazer & Coulter, 2012; Bannink, 2012). Clients are supported to see their depression as a normal part of life and something they can proactively manage, meaning they are able to live a life they wish to lead even alongside periods of low mood (logic model box 5). Therapists utilise techniques to enhance memorability and generalisability of sessions throughout (attention recruitment, categorisation, evaluation, application, repetition, practising remembering, use of cue based reminders, and praising recall; see Harvey et al., 2014). Clients are also given handouts summarising key therapy content to ensure long term recall of session content (logic model box 6).

Early Sessions

Session 1 assesses depression and presents the ADepT rationale, in particular that the treatment aims to build wellbeing in important life domains and sees depression as a barrier to achieving this goal. Clients are given a simplified version of the logic model underpinning the intervention, which the therapist tailors and personalises to them. Sessions 2 to 5 help clients establish what is important to them in life (their

values) in vocation, relationship, hobbies and self-care domains; set behavioural goals consistent with these values; and breaks these goals down into 'action-steps' (logic model box 2).

A key element introduced in this early phase values work is a modified values 'bullseye' tool that is co-developed with the client (cf. Lungdren et al., 2010). In later sessions, clients report if they are completing the action steps they specified and if they are moving closer to the bullseye in each life domain. Values, goals and action-steps are updated in an ongoing fashion as therapy proceeds. This tool serves many of the functions of a conventional therapy formulation (synthesis of client experience with theory; normalisation of presenting issues; promoting client engagement; making goals seem manageable; guiding selection, focus and sequence of treatment; identifying strengths and building resilience; monitoring outcomes; guiding supervision; see Kuyken, Padesky & Dudley, 2010) without having any emphasis on psychopathology. This values work can either involve reconnecting clients to previous values they have neglected while depressed and/or helping clients build new values in areas of life they have not considered before. Often there is an emphasis on distinguishing between what is intrinsically important to clients, versus what values may have emerged as a reaction to depression or external pressures.

Middle Sessions

Sessions 5 -12 behaviourally activate clients to work towards the goals set on the values tool (logic model box 3). To maximise the likelihood that these goals are completed, clients are encouraged to recognise if they have the capability, opportunity and motivation to carry out each action step (known to be key underlying determinants of behaviour change; Michie, Stralen & West, 2011). If they lack any of these determinants, coaching approaches are used to build them in session. As they work towards these goals, clients are supported to understand how depressogenic patterns of thinking and behaviour hinder them from fully embracing opportunities to enhance PA (thriving; maximising PVS activation) and coping with challenges to minimise NA (being resilient; minimising NVS activation) (boxes 4a and 4b in the logic model). The therapist supports the client to recognise and then 'act-opposite' to a range of depressogenic mechanisms including: avoidance, rumination, self-criticism, negative processing biases, and neglect of sensory experience.

A mapping tool is introduced to facilitate this process of pattern recognition and modification (a modified 'hot cross bun'; cf. Padesky & Mooney, 2012) in the middle phase of therapy. Clients are encouraged to map out a 'depressed me' (how depression makes them cope), then develop an 'alternative me' (alternative ways of coping to maximise chances of goal attainment) and commit to behavioural experiments to test out the 'alternative me' (cf. Padesky & Mooney, 2012). This can involve reconnecting clients to their existing strengths (cf. Gielen, van Woerkom & Meyers, 2017) or developing new skills as appropriate. This mapping occurs at the 'micro' rather than 'macro' level. The emphasis is on understanding the fine details of what helped or hindered in a specific context rather than at an abstract, general level. The focus of these maps is the present and future, with far less attention paid to historical insights than in mainstream CT. Therapists aim to model and actively reinforce a (realistically) positive and future-oriented cognitive and interpersonal style (MacLeod, 2017; Vilhauer, 2014; Verplanken, 2006; Wood & Neal, 2007).

Therapists use a mixture of cognitive change, role-play, skills training and problem-solving techniques to help clients build new positive patterns of response. Cognitive change work predominantly adopts a 'utility' rather than 'truth' perspective (thinking/attending in way that maximises the chances a goal is achieved rather than evaluating whether particular ways of thinking/attending are realistic). Other tools used in this middle phase of therapy include a positive journal (where clients record moments of resilience and thriving occurring in their everyday life) and brief training in mindful engagement with the external senses during everyday activity (both as a way to enjoy every day 'simple pleasures' and as a way to regulate negative mood at times of stress).

End and Booster Sessions

Sessions 13-15 develop a wellbeing plan, to help clients continue to move towards positive recovery. Therapy is framed as the start of this process. The intention is for clients to consolidate skills they have learned and build a habit of maintaining and tracking progress towards valued goals. By doing so, in the event of a future drop in mood clients will have have well-rehearsed, and increasingly automated, positive coping strategies to deploy (box 6 of logic model). The wellbeing plan involves working through a series of handouts consisting of: a review of progress made in therapy (including celebrating progress); setting goals

for the future on the bullseye tool; considering what could help or hinder achieving these goals; and developing concrete steps outlining how to respond to any signs of relapse or mood lift. Sessions 16-20 are optional booster sessions, focusing on managing any mood dips and reviewing and updating the wellbeing plan and values bullseye tool as required.

Similarities and Differences with Other Approaches

The systematic emphasis on thriving (making the most of opportunities to enhance positive affect and PVS function) distinguishes ADepT from CBT and MBCT (which predominantly focus on NVS regulation). ADepT is similar to ACT in that it focuses on the pursuit of valued activities. However, the ACT stance is to engage in these activities regardless of unwanted experiences that may occur and with little attempt to alter symptoms. In contrast, ADepT explicitly focuses on trying to enhance wellbeing and PA and reduce NA and symptoms whilst engaging in these valued activities. ADepT also has overlap with Behavioural Activation approaches (Martell et al., 2010), but with more of an explicit emphasis on clarifying values and targeting cognitive mechanisms that may sabotage resilience and thriving during planned activities.

Therapists and Supervision

Treatment was delivered by three therapists. Therapist one was the first author of this study and the intervention developer (a doctoral level clinical psychologist; BD). Therapist two was a doctoral level clinical psychologist. Therapist three was a BABCP accredited nurse therapist. Group supervision was provided weekly for 60m in group format (led by BD).

Measures

Diagnostic Interview

The Structured Clinical Interview for DSM-IV-TR Axis I Disorders (SCID-I; First et al., 1994) was used to assess whether participants met criteria for a Current Major Depressive Episode and to rule out current or past mania or psychosis at intake. At one-year follow-up, participants also underwent a diagnostic interview to assess if they had met criteria for a major depressive episode in the period of time since

completing acute treatment using The Longitudinal Interval Follow-up Evaluation (LIFE, Keller et al., 1987). The interviewer was not blind to treatment phase.

Weekly Measures

The 14-item Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS; Tennant et al., 2007) was used to index wellbeing and the 9-item Patient Health Questionnaire (PHQ-9; Kroenke et al., 2001) was used to measure depression symptom severity. Both measures have been shown to be sensitive to change during treatment (Maheswaran et al., 2012; Kroenke et al., 2010) and had acceptable reliability in the current sample (see supplementary materials Section S1 for full descriptions, reliability co-efficients, and normative data). These measures were filled in weekly, either via post or online.

Acceptability

After the acute treatment phase was completed, participants were invited to take part in a qualitative interview (lasting approximately 45 minutes) to explore their experience of ADepT treatment. A particular focus of the qualitative interviews was to examine if the wellbeing, solution-focused emphasis of ADepT would be feasible and acceptable to clients, given concerns that such an approach could be experienced as "PollyAnna-ish" (see Dunn, 2012). A pragmatic stance was adopted (stimulating the combination of action and reflection to solve 'real world problems'; Feilzer, 2010; Biesta, 2010). The interviews were conducted by the research team and the qualitative analysis was completed by the first author (BD). The interviews followed a topic guide to ensure that all areas were covered, but this was used flexibly so as to allow other issues of importance to participants to be fully examined (see Supplementary Materials Section S4). At the end of this interview, participants also quantitatively rated the acceptability of the intervention (from 1 = not at all acceptable to 5 = extremely acceptable); how satisfied they were with the intervention (from 1 = not at all satisfied to 5 = extremely satisfied); and whether they would recommend this treatment to friends or family suffering from depression (from 1 = not at all likely to 5 = very likely). Qualitative interviews were audio-recorded and then transcribed. Thematic analysis of the interviews was conducted using a Framework approach, involving the coding and sorting of textual units according to both deductive and inductively-derived categories, and the use of matrices to review the coded data, investigate commonalities and

differences and search for patterns (Ritchie, Spencer, & O'Connell, 2003) (Gale, Heath, Cameron, Rashid, & Redwood, 2013).

Extended assessment measures

At intake, pre-treatment, post-treatment, two-month follow-up and approximately one-year follow-up participants completed an extended battery of measures. In addition to the WEMWBS and PHQ-9, participants completed four further measures. The 5-item Work and Social Adjustment Scale (WSAS; Mundt et al., 2002) was used to assess functional impairment resulting from depression. The past week version of Positive and Negative Affect Scale (PANAS; Watson et al., 1988) was used to measure levels of positive and negative affect (measures of PVS and NVS respectively). The 7-item Generalized Anxiety Disorder scale (GAD-7; Spitzer et al., 2006) was used to measure anxiety symptom severity, given depression is frequently comorbid with anxiety. The 14-item Snaith Hamilton Pleasure Scale (SHAPS; Snaith et al., 1995) was to measure anhedonia severity, reporting the continuous scoring recommended by Franken et al. (2007). Reliability was acceptable for all measures (see supplementary materials Section S1 for full descriptions and reliability coefficients).These measures were either administered via post or in a face-to-face interview conducted by a member of the research team (EW)[2] . The mean interval from final acute therapy session to one-year follow-up was 406 days (SD=92.42, range = 241 days to 492 days).

Results

Recruitment, data completeness and demographic characteristics

Figure 2 presents the consort diagram for the case series. Fourteen participants entered the case series but three were excluded (P1 and P2 as they recovered during the baseline phase [PHQ-9 change > 6 points and scoring < 10 at pre-treatment]; P14 because it became apparent their primary presenting problem was a head injury[3]). All remaining 11 participants had complete PHQ-9 and WEMWBS data for each weekly assessment, apart from P10 (no treatment week seven data) and P13 (no baseline week two data). We substituted these two missing values with the average of the two values either side of them. Complete extended assessment data at pre, post and two-month follow-up were available for ten out of eleven participants (P4 did not complete post-treatment and follow-up assessments). We substituted the post-treatment PHQ-9 and WEMWBS scores for P4 with weekly measures taken at session 15. Nine out of

eleven participants took part in the post-treatment qualitative interview. Eight out of eleven participants were contactable after one-year for the additional follow-up, all of whom agreed to take part. P4, P9, P13 could not be contacted at one-year follow-up, all of whom had showed reliable improvement on the PHQ-9 during acute treatment.

Table 1 summarises client demographic and clinical characteristics at intake. The sample were aged between 20 and 70, were predominantly female, were mostly not employed, and all were of White British ethnic origin. None declared themselves as having a disability. Mean depression severity was in the moderately severe range, mean anxiety severity was in the moderate range, and mean WSAS severity was in the severe range. Mean WEMWBS score was well below general population averages (mean for 7020 UK adults from Health Survey for England 2011 data = 51.61; SD=8.71); mean SHAPS score was well above general population averages (mean for 50 members from the Dutch general population=20.2, SD=4.4; Franken et al., 2007); and mean PANAS PA was well below, and PANAS NA was well above, general population averages (mean for 2,527 general adult population: PA = 31.72; SD= 7.38; NA = 17.04, SD=6.68; Crawford et al., 2009). The sample predominantly had recurrent depression (mean of 8 episodes, with mean age of first onset in their twenties), with three participants having attempted suicide in the past at least once. Eight out of eleven were taking anti-depressants, none of whom reported a change in medication during the baseline phase and three of whom reported a change during acute treatment. One patient switched medications and two reduced or stopped medications, meaning this is unlikely to have inflated treatment response (unless the medications were serving an iatrogenic function). Ten out of eleven had undergone previous psychological treatment (a combination of counselling, high intensity CBT, low intensity CBT, low intensity BA, and MBCT), none of which overlapped with treatment in the present study. Therapist one treated three cases and therapists two and three each treated four cases.

Acceptability and feasibility of ADepT

Detailed thematic analyses of the qualitative interviews is presented in the supplementary materials (Section S4). Here we summarise key findings from this analysis. Nine out of eleven clients perceived ADepT to have been very helpful, finding the future-oriented wellbeing focus unexpected but beneficial and leading to improvements in depression. One client saw depression as a secondary consequence of

fibromyalgia, which they did not feel the therapy had addressed. Another client wanted concrete advice about a life issue, and was disappointed the therapist explored this Socratically rather than providing answers. A number of clients reported developing a new relationship to depression during ADepT, seeing it as just one part of them that they could learn to control to maximise their resilience (i.e leading to meaningful changes in self-identity). Aspects of ADepT that were viewed as particularly helpful included: the establishment of a positive, caring therapeutic relationship; identifying and living according to their values; building self-care; engaging with simple everyday pleasures; socially reconnecting; having clear goals; and having a practical focus. The handouts and exercises were viewed as constructive, although the volume of homework felt overwhelming to some clients (in clients where English was not the first language or who had dyslexia). Booster sessions were seen as useful, even though not all clients felt the need to take them up. Areas to improve in the ADepT protocol identified by clients included redrafting some handouts to be clearer, further emphasising social reconnection, and taking life stage into account more explicitly. Barriers to engagement included negative previous experiences of talking therapy, depression reducing motivation, and practical life issues. The three clients whose depression had not remitted during acute treatment all gave positive feedback about the treatment, identifying clear behavioural gains they had made despite only partial or no symptomatic relief.

Table 2 reports number of sessions attended in the acute treatment phase and optional booster phase and participants quantitative' ratings of ADepT. Indicating ADepT was acceptable to clients, four were extremely satisfied and five were very satisfied with treatment. Four rated the treatment as extremely acceptable and five rated the treatment as very acceptable. Eight participants would be extremely likely to recommend ADepT and one would be likely to recommend ADepT. Further suggesting ADepT is acceptable, ten out of eleven participants completed the full dose of 15 acute sessions. The other participant chose to stop treatment after eleven sessions, having made a full recovery (PHQ-9=0 and GAD-7=2 at final treatment session) and met their treatment goals. A minimum adequate dose was defined as attending at least 50% of sessions and all participants exceeded this criteria.

Eight out of eleven participants made use of the ADepT optional booster sessions (attending on average 2 sessions, SD=1.94). Of those participants who engaged with boosters but did not complete all

five, all agreed with their therapist at which point to terminate the booster phase (typically at the point they had met their treatment goals and/or stayed well for an extended period of time). Of those who did not engage with any booster sessions, two participants declined as they felt they were currently well and had no need of further sessions at the end of acute treatment (both with PHQ-9 score < 10 at post treatment). One participant initially accepted the offer of booster sessions but then withdrew prior to the first scheduled booster as they felt they were well and had no need of further sessions.

Time series analyses of weekly depression and wellbeing data

Simulation-modelling analysis (SMA; Borckardt, 2006) was used to examine statistically the baseline and acute treatment time series data for PHQ-9 and WEMWBS, whilst taking into account the auto-correlation inherent in time series data of this kind (results summarised in Table 3). We focused on both level and slope of change (Kratochwill et al., 2010). Significant differences in level and/or slope indicate an intervention effect for each individual participant. Given that therapeutic interventions for depression are lengthy and rarely result in an immediate improvement in symptoms (especially in treatment resistant cases), an intervention effect is more likely to be observed in the slope analysis than the level analysis. Level change analysis examines if there is a difference in mean level during the baseline and intervention phase. Slope change analysis examines the fit of the data to an *a priori* specified model of a flat slope during baseline and then a linear improvement during acute treatment. SMA generates Pearson's r correlation coefficients to indicate the extent of fit with these models. Multiple simulations of data streams sharing the same characteristics as the observed data (same degree of autocorrelation and same length of the two phases) are then run to determine the probability of obtaining the observed correlation with the *a priori* level-change and slope-change (using conventional p value cut-offs to indicate significance). We did not include the intake assessment point in the time series analyses, as for some participants there was more than a week gap between intake assessment and the start of the baseline period.

Autocorrelation between time measures (using a phase lag of one) varied between -.07 and .71 for PHQ-9 and between .08 and .87 for WEMWBS. No participants showed a significant difference in mean level of either outcome in the treatment relative to the baseline phase. For both PHQ-9 and WEMWBS, the fit of the slope change model was significantly greater than would have been expected by chance for seven

out of eleven participants (P4, P6, P7, P8, P9, P11, P13), indicating significant linear improvement over time. The supplementary materials Section S2 reports visual analysis of these time series data, with Supplementary Figure 1 plotting weekly levels of depression and wellbeing for each participant. This visual analysis reached broadly conclusions to the SMA analysis, with the same seven participants appearing to improve on both outcomes. However, the pattern of change was not clearly linear, with significant variability over time within and between participants. This likely reflects the fact that mid-treatment clients were moving from a position of avoidance to interacting with potential opportunities and challenges in life. Further, clients were refining new skills to manage these opportunities and challenges, making it very likely they would experience more variability in mood in this part of treatment. According to case series guidelines, to demonstrate an intervention effect there needs to be either a significant change in slope *or* level (Kratochwill et al., 2010), which is replicated at least three times. This analysis meets this criterion, although is only considered moderate evidence as it was not demonstrated for each individual case.

Reliable and clinically significant change analyses

Following Jacobson & Truax (1991), we looked at the proportion of individuals showing reliable change (RC) and reliable and clinically significant change (R+CSC) at each follow-up period, relative to pre-treatment levels. We focused on both improvement and deterioration to identify benefits and harm. Reliable change was computed for each measure by dividing the magnitude of change observed during the course of therapy by the standard error of the difference score. Clinically significant change analyses used criterion c (participants are closer to comparison than clinical mean) for all measures apart from the WSAS. As no healthy comparison group values were available for the WSAS, criterion a (falling more than two standard deviations away from the clinical group mean in the direction of clinical improvement) was instead deployed for this measure. The Leeds Reliable Change Indicator Excel tool (Morley & Dowzer, 2014) was used to conduct these analyses. This requires the user to enter reliability estimates and mean (SD) values for a clinical and comparison sample. The pre-treatment assessment point in the present sample was used to define clinical group values for each measure. Reliability estimates (test-retest reliability where available and otherwise internal reliability) and comparison sample values were derived from relevant scale validation papers (see Supplementary Materials Section S1 for full details).

Table 4 summarises the results of these analyses. For the PHQ-9, RC was observed for eight out of eleven participants (73%) and R+CSC was observed for seven out of eleven participants (64%) at the post-treatment assessment point. For the WEMWBS, RC was seen for seven out eleven participants (64%) and R+CSC improvement was seen for six out of eleven participants (55%) at the post-treatment assessment point. These improvements were preserved for a majority of participants at each subsequent follow-up point for both primary outcome measures (RC>50%; R+CSC>38%).

There were also gains for a majority of participants at all time points for WSAS functioning (RC≥50%, R+CSC≥40%) and PANAS positive and negative affect (RC≥60%, R+CSC≥50%). GAD-7 improvements at post and two-month follow-up assessment observed for a smaller proportion of participants (RC≥30%; R+CSC≥30%) but there were gains for a majority of participants at one-year follow-up (RC=75%; R+CSC=75%). Only a minority of participants showed clear improvement for SHAPS anhedonia at each time point (RC≥30%; R+CSC≥20%). There was little evidence of reliable deterioration except that P5 deteriorated on the WEMWBS at two-month follow-up and PANAS positive affect at two-month and one-year follow-up. P7 also deteriorated at one-year follow-up on PANAS negative affect.

Recovery Analyses

Next we examined the proportion of clients showing clinical levels of symptoms on the PHQ-9 and GAD-7 at each time point. According to recent meta-analyses, a PHQ-9 score of 10 has 85% sensitivity and 89% specificity for diagnosing major depressive disorder (Manea, Gilbody & McMillan, 2012) and a GAD-7 score of 8 has 83% sensitivity and 84% specificity for diagnosing generalized anxiety disorder (Plummer et al., 2016). Scoring beneath these cut-off values is used to indicate clinical ‘recovery’ in UK IAPT services (from where the present sample was predominantly recruited). At intake and pre-treatment all eleven participants (100%) reported clinically significant levels of depression and ten out of eleven (91%) participants reported clinically significant levels of anxiety. Depression recovery criteria were met by eight out of eleven (73%) participants at the post-assessment, five out of ten (50%) participants at the two-month follow-up assessment, and four out of eight (50%) participants at one-year follow-up assessment. Anxiety recovery criteria were met by five out of eleven (45%) participants at post-assessment, five out of ten (50%) participants at two-month follow-up, and five out of eight (63%) participants at one-year follow-up. UK

IAPT services additionally report the proportion of clients showing reliable improvement, defined as at least a six point drop on the PHQ-9 and at least a four point drop on the GAD-7 during acute treatment. Seven out of eleven participants (64%) met criteria for reliable improvement on the PHQ-9 and eight out of eleven participants (73%) met criteria for reliable improvement on the GAD-7.

The diagnostic interview at one-year revealed that six out of eight participants [75%] interviewed had not relapsed during the follow-up period and therefore met criteria for sustained recovery. Of the two participants who had met diagnostic criteria during the follow-up period, both of these had scored above the PHQ-9 caseness cut-off immediately post treatment (PHQ-9 post-treatment was 14 for P3 and 25 for P5), indicating they had not fully responded to acute treatment. P5 met criteria for a current major depressive episode at one-year follow-up on the SCID and continued to score above the PHQ-9 cut-off. P3 was no longer in episode according to the SCID (although her PHQ-9 score was 10, so just above the caseness cut-off).

Group level analysis

Figure 3 plots mean levels for each outcome measure at each assessment period. Total raw scale scores are plotted for measures, with the exception of the PANAS which has been converted into Z-scores (relative to general population normative data) to aid interpretation of the magnitude of PA versus NA change. General population average levels and cut-offs for clinical caseness are super-imposed on the Figures to help interpret the extent to which intake deficits are 'normalised' during treatment.

We conduct complete case analysis, rather than intention to treat analysis, as it was not possible to conduct multiple imputation (the model would not converge; see Mcneish, 2017). There is a balance between controlling for type I and type II error in small sample research of this kind, with use of non-parametric approaches and correction for multiple comparisons potentially inflating type II error, whereas use of parametric approaches and non-correction for multiple comparisons potentially inflates type I error. We chose to use parametric analysis, as inspection of the paired differences scores suggested an approximately normal distribution of change scores. As we had clear *a priori* hypotheses about each measure, we chose not to control for multiple comparisons. Instead, we restricted the number of comparisons being run by first conducting a series of repeated measures ANOVA on each outcome variable

(with time as the within-subjects factor: intake, pre, post and two-month follow-up) and only proceeding to pairwise comparisons in cases where the ANOVA for that measure was significant. We did not include the one-year follow-up outcomes in the repeated measures ANOVAs as this assessment was a post-hoc amendment to the protocol and data were only available for eight participants. Including the one-year follow-up data in the ANOVA would have significantly reduced statistical power.

As intended, there was a main effect of time for PHQ-9, $F(3,27)=20.32$, $p<.001$, $\eta^2_p=.69$, WEMWBS, $F(3,27)=11.82$, $p<.001$, $\eta^2_p=.58$, WSAS, $F(3,24)=11.71$, $p<.001$, $\eta^2_p=.59$, GAD-7, $F(3,27)=10.61$, $p<.001$, $\eta^2_p=.54$, PANAS PA$(3,27)=8.56$, $p<.001$, $\eta^2_p=.49$, PANAS NA, $F(3,27)=7.97$, $p=.001$, $\eta^2_p=.47$, and SHAPS, $F(3,27)=8.68$, $p<.001$, $\eta^2_p=.49$.

We resolved these main effects of time by running pairwise comparisons for each measure using paired sample t-tests. As intended, there was no significant difference for any measure from the intake to pre assessment (indicating a stable baseline phase), ps>.12.There was also significant improvement intake to post, ps<.005, intake to two-month follow-up, ps<.03, pre to post, ps<.02, and pre to two-month follow-up, ps<.05. The one exception was that the pre to two-month follow-up comparison for PANAS PA was not significant, p=.16. There was no significant change from post to two-month follow-up for any measures, ps>.10, with the exception of improvements in PANAS PA over that time period, p=.01 (indicating benefits were sustained or extended during the follow-up period).

We then went on to look at the differences between each of the earlier assessment points and the one-year follow-up assessment on the subset of eight individuals who had complete data with paired sample t-tests. As intended, one-year follow-up levels were significantly improved relative to intake and pre-treatment, ps<.03, but did not differ relative to post-treatment or two-month follow-up, ps>.13, for PHQ-9, WEMWBS, PANAS PA, and WSAS. A similar pattern emerged for the SHAPS and GAD-7, with the exception that symptom levels had further significantly improved from two-month follow-up to one-year follow-up, ps<.04. For PANAS NA, one-year follow-up levels were only trend lower than pre-treatment NA levels, p=.054, but otherwise the pattern was the same as for PHQ-9, WEMWBS, PANAS PA and WSAS.

Benchmarking analysis

We benchmarked the group level case series outcomes against effect sizes from routine practice collected in the same service setting and high quality RCTs of best practice that had included the same measures of PVS and NVS function. We selected the CPT2 trial (16 week outcomes for depression, PA and NA in the CT arm; DeRubeis et al.,2005), the CPT3 trial (six month outcomes for depression, PA, NA in the combined CT and anti-depressant medication arm; Hollon et al., 2014), the COBRA trial (six month PHQ-9 depression GAD-7 and anxiety outcomes following CBT and BA at six months; Richards et al., 2016), the COBALT trial (six month PHQ-9 depression and GAD-7 anxiety following CBT at six months; Wiles et al., 2013), and routine IAPT outcomes (pre-post outcomes for PHQ-9 depression, GAD-7 anxiety and WSAS functioning for combined high and low intensity treatment; Zahtra et al., 2015; pre-post outcomes for WEMWBS wellbeing for high intensity treatment; Widnall et al., submitted) for benchmarking.

Figure 4 plots the pre to post and pre to one-year effect sizes on each outcome measure in the case-series against outcomes from these other datasets. For all measures at both time points, effect sizes following ADepT were large (Hedges g, correcting for small sample sizes, >0.8). Reflecting the sample size in the case series, the 95% confidence intervals of the effects were broad and crossed zero. Visual inspection of mean effect sizes tentatively suggests ADepT outcomes were comparable to existing trials for depression, anxiety, and NVS outcomes, but potentially superior to existing trials for wellbeing, anhedonia and PVS outcomes. ADepT outcomes also appeared superior to routine IAPT service outcomes for anxiety, depression, and functional impairment.

Discussion

This multiple randomized baseline case series evaluated the acceptability and clinical efficacy of Augmented Depression Therapy (ADepT), a novel psychological treatment that aims to target positive valence system (PVS) and negative valence system (NVS) deficits simultaneously so as to both reduce symptoms of psychopathology and build wellbeing in depressed clients.

Participants were willing to enter the case series and there was a high degree of engagement with ADepT. All participants completed acute treatment with a planned discharge and a majority of participants attended some booster sessions. Given high dropout rates in individual CBT for depression (estimated to be greater than 25%; Hans & Willer, 2013; Fernandez, Salem, Swift & Ramtahal, 2015), the retention rates in

the case series are encouraging. Participants rated ADepT as acceptable, that they were satisfied with it, and would recommend it to friends and family. The qualitative interviews revealed that the wellbeing focus of ADepT was acceptable to and valued by clients, with no evidence that it was experienced as "PollyAnna-ish" (cf. Dunn, 2012).

There was preliminary but nevertheless promising evidence of clinical effectiveness on the primary outcome measures, with triangulation in findings across analytic methods. At the individual level, time series analyses showed significant improvement on depression and wellbeing for 7/11 participants. Reliable improvement was observed for 8/11 participants on depression and 7/11 participants on wellbeing (9/11 participants improving on at least one of these measures). There was also reliable improvement for most participants on secondary outcome measures of functional impairment, anhedonia, PA and NA. Recovery analyses found that 7/11 participants scored beneath clinical cut-offs for depression at post-treatment. Group level analyses found significant improvements on all outcome measures, all of a large effect size (albeit with broad confidence intervals due to the small sample size). Gains were mostly preserved (and for wellbeing and anhedonia enhanced) at follow-up.

Benchmarking analyses indicated that depression and anxiety mean effect-sizes in the case series were broadly comparable to those observed in existing CBT and BA trials (e.g., Richards et al., 2016; DeRubeis et al., 2005; Hollon et al., 2014) and were superior to routine IAPT care (e.g., Zahtra et al., 2015). Mean effect sizes for ADepT for PVS outcomes (anhedonia and positive affect) were larger than those found in previous trials of BA and CBT (DeRubeis et al., 2005; Hollon et al., 2014; Richards, 2016), suggesting ADepT has potential to lead to step-wise improvements in the capacity to treat PVS deficits. Importantly these bench-marking analyses take into account a range of non-specific factors (expectancy effects from being involved in research, benefits of a positive therapeutic alliance, spontaneous recovery) that could be biasing interpretation, as these comparison datasets are also subject to these non-specific effects. Therefore, where ADepT shows a larger effect size than these other treatments, it is plausible this is specifically related to treatment. Again, the small sample size (and as a result broad confidence intervals around the effects reported), mean these bench-marking results need to be interpreted cautiously.

It is important to examine if treatment response varies as a function of intake depression severity. A significant pre-post improvement for WEMWBS and PHQ-9 in time series analysis was shown for four out of five participants in the mild PHQ-9 range (10-14); for two of three participants in the moderately severe range PHQ-9 range (15-19); and for one out of three participants in the severe range (>19). While further evaluation is required, this suggests better responses at the less severe end of the sample in the current version of the protocol and that some adaptations to the protocol may be required for severely depressed clients.

While outcomes were generally promising, two participants did not respond optimally despite engaging fully with the acute treatment sessions. P3 failed to show reliable improvement in depression and wellbeing immediately post-treatment, but did engage with some of the booster sessions and showed significant gains on both outcomes at one-year-follow-up. P5 also failed to show reliable improvement for either primary outcome measure immediately post-treatment. Despite attending all of the booster sessions, this client reported no further improvement in wellbeing or depression at subsequent follow-ups (and showed deterioration on wellbeing at two-month follow-up). Despite these mixed outcomes, both clients rated the treatment as acceptable and in qualitative interviews reported having made meaningful behavioural changes as a result of engaging in sessions.

The sample were predominantly cases of moderate to severe, treatment resistant depression, with comorbid anxiety, recruited from NHS waiting lists. In particular, most had not responded to prior anti-depressant or other psychological therapies delivered as part of their routine care. They are therefore representative of patients seen in routine NHS care (e.g. Zahtra et al., 2015) and comparable to those treated in existing trials that we used for benchmarking purposes (Richards, 2016; DeRubeis et al., 2005; Hollon et al., 2014). There were no obvious differences in outcomes between therapists, preliminarily suggesting effects are not carried by particular therapists and that the protocol is likely generalizable to different workforces.

The case series met the pre-specified continuation rules in full. Nine out of eleven participants (82%) showed reliable improvement from pre to post treatment on the PHQ-9 and/or WEMWBS (82%; exceeding the 60% target). No clients showed clinically significant deterioration pre to post treatment on either the

PHQ-9 or WEMWBS (0%; falling beneath the 30% target). All eleven clients completed a minimum adequate dose of treatment (100%; exceeding the 50% target). 161 out of a possible 165 acute treatment sessions offered across clients were attended (98%; exceeding the 60% target). All participants rated ADepT treatment as acceptable, were satisfied with ADepT, and would recommend ADepT to friends and family (100%; exceeding 60% target). There were no serious adverse consequences for participants that were clearly trial or treatment related. Pre-post and effect sizes for depression and wellbeing were both large in magnitude (Hedges $g > .08$). On this basis, it seems valid to proceed to further evaluation of ADepT.

The bespoke continuation rules used in the present case series were *a priori* set to be liberal, as this was the first iteration of the ADepT protocol and the treatment was not expected to be fully optimised at this stage. In later stages of evaluation of subsequent iterations of ADepT, more stringent continuation rules will be applied (whereby ADepT will need to be at least as effective as current practice in a pilot trial before proceeding to definitive trial). It is noteworthy that the present case series results clearly exceeded these liberal continuation rules, suggesting ADepT has potential to meet more stringent continuation rules in later development phases.

It is helpful to consider whether ADepT is likely to add value to existing depression treatment options, given that there is already a plethora of depression therapies all with broadly equal (but sub-optimal) efficacy. To add value new therapies need to strike a balance between innovation and building on what has come before. A position of extreme novelty, where a treatment discards attempts to repair thinking, behaviour and affect using proven intervention elements from existing treatments, is unlikely to be optimally effective, acceptable to key stakeholders, and implementable (as it will not capitalise on existing skills in the workforce). However, a therapy that simply renames and repackages existing treatment elements without any novelty is unlikely to lead to sufficient improvements in outcome to warrant extensive investment. ADepT attempts to find a position of optimal novelty by enhancing existing treatment elements from CBT, ACT and mindfulness approaches to target PVS and NVS disturbances (‘horizontal innovation’) and ensuring these are deployed to target wellbeing and positive recovery as well as symptom relief (living well alongside mental illness; ‘vertical innovation’). Shifting the emphasis to enhancing wellbeing and functioning has potential to represent a radically trans-diagnostic treatment approach.

Ultimately, establishing the value in ADepT is an empirical question. Existing treatments like BA and CBT are relatively ineffective at repairing PVS disturbance and broader wellbeing deficits (Dunn et al, submitted; Dunn et al, in prep). The present case series suggests ADepT has potential to lead to stepwise gains in PVS and wellbeing. While other treatments are also being developed to target PVS disturbance with promising preliminary results (Chaves et al., 2017; Craske et al., 2019; Geschwind et al., 2019; Ruini & Fava, 2012; Taylor, Lyubomirsky & Stein, 2017; see Chakhssi et al., 2018), ADepT in our view is unique in its joint focus on both PVS *and* NVS regulation. RCTs now need to be conducted that compare ADepT against other treatments.

A key part of the ADepT rationale is that it aims to build wellbeing *and* reduce depression, in part based on the logic that these are orthogonal constructs (Keyes, 2005). However, there was considerable symmetry in the pattern of response across the PHQ-9 measuring depression and the WEMWBS measuring wellbeing in the present study (i.e. participants either improved on both or neither measure). Inspection of the individual items of the WEMWBS and PHQ-9 indicates there is substantial overlap in their content (although clear difference in whether items are positively or negatively phrased). This echoes recent psychometric findings that shows the WEMWBS is not clearly distinct from symptom measures (Bohnke & Croudace, 2016). It may be that the PHQ-9 and the WEMWBS are not ideal measures of symptoms and positive recovery respectively and there is a need to develop better tools to measure these orthogonal constructs. Consistent with this position, in the qualitative interviews, some clients described improvements in values-consistent functioning that were not captured by either measure (for example, the two non-responders both described behavioural gains they had made during treatment). Alternatively, it may be that wellbeing and depression are better viewed as opposite ends of a single underlying bipolar dimension (running from depression to neutral to wellbeing). Even if this is the case, in our view there is still merit in measuring wellbeing outcomes given clear evidence of the benefits of setting approach rather than avoidance goals to bring about behavioural change (Elliot et al., 1997). Further, clients definitions of recovery from depression focus more on desired positive outcomes than symptom relief (Zimmerman et al., 2006) and participants in the case series reported particularly valuing the wellbeing focus of ADepT.

There are a number of limitations that need to be held in mind regarding the present findings. First, the lead author was treatment developer, supervisor and one of the therapists, meaning that there is potential for allegiance biases. Second, treatment fidelity was not assessed (as no fidelity tool currently exists). Third, the time series slope analyses fitted a linear model of change during treatment (followed SMA guidance; Borckardt, 2008), although visual inspection of the data suggested a non-linear, variable pattern of change. The effect of this non-linearity would go against our experimental predictions (of a significant intervention effect), so will have inflated type II rather than type I error. Fourth, the small sample means that the effect sizes have wide confidence intervals and are likely to be positively biased (although Hedges small sample correction was used to minimize this). Fifth, the group level analyses were under-powered (as we did not recruit the intended completer sample size of 13) and we did not correct for multiple comparisons as a result (to balance the risk between type I and type II error). It is noteworthy that across the board the group level analyses were all clearly significant for key comparisons, meaning it is very unlikely that this pattern would have emerged solely due to random error or inflated type I error. Further, the time series analyses on the primary outcome measures were adequately powered and were systematically replicated across a number of participants. Sixth, DSM-IV rather than DSM-V instrumentation was used to assess depression. While no major changes were made to diagnostic criteria for depression in the shift from DSM-IV to DSM-V, the bereavement exclusion was removed and additional specifiers of 'with anxious distress' and 'with mixed features' were added. We did not exclude any eligible participants on the basis of the bereavement criteria in DSM-IV, meaning that our diagnoses align with those that would have been reached under DSM-V. However, we did not capture the additional specifiers. Seventh, the Caucasian, predominantly middle aged sample recruited from a single site in the South West of England lacked diversity. It remains to be established if these effects will generalise to other populations. Eighth, while the post treatment quantitative and qualitative data suggested ADepT was acceptable and feasible, we did not capture participants'views of the approach after their first treatment session (for example, using the Credibility/Expectancy Questionnaire; Devilly & Borkovec, 2000).Finally, we did not pre-register the case series protocol (although we report data on all outcome measures collected using a variety of analytic procedures, therefore minimising risk of p- or

measure- hacking). These issues will be partly addressed by an ongoing pilot randomized trial comparison of ADepT versus CBT (Dunn et al., 2019).

In summary, this case series provides preliminary evidence that ADepT is an acceptable intervention to clients that is likely to lead to improvements in PVS and NVS and by doing so lead to enhanced wellbeing and reduced depression symptoms. The case series met all the continuation rules, indicating ADepT is ready to proceed to more rigorous evaluation via randomized controlled trial methodology.

Acknowledgements

The present work is funded by a Career Development Fellowship awarded to the principal investigator (BD) by the National Institute for Health Research in the UK, grant number CDF-2014-07-10. The views expressed in this protocol are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health. The AccEPT service that hosted this research is commissioned by the National Health Service NEW Devon Clinical Commissioning Group. Thanks to Devon Partnership Trust for assistance recruiting patients; to Pete Mason and Jo Mackenzie for input into development, delivery and refinement of the ADepT protocol; Sarah Cooling for assistance with qualitative interviews; David Richards, John Powell, Nicola Wiles, Rob DeRubeis and Steve Hollon for permission to use their trial data for benchmarking analyses; and Rod Taylor, Anne Spencer, Andy McLeod, Richard Moore, Gerjo Kok, Gerda Kraag, Michelle Moulds, Rob DeRubeis, David Richards, Kim Wright, Rachel Handley, David Jeffery, and the Exeter Lived Experience Group for input into the broader ADepT project. Thanks to Steve Kellet, Stephen Barton, J.D Smith, and Gary Brown for feedback on an earlier draft of this manuscript and to Sophie Dunn and Laura Warbrick for proof reading.

References

Al-Janabi, H., Flynn, T. N., & Coast, J. (2012). Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. *Quality of Life Research*, *21*(1), 167-176.

Andrews, G., & Titov, N. (2007) Depression is very disabling. *Lancet*, *370*, 808-809.

Anthony, W. A. (2003). Expanding the evidence base in an era of recovery. *Psychiatric Rehabilitation Journal*, 27, 1-2.

Bardeen, J. R., Tull, M. T., Stevens, E. N., & Gratz, K. L. (2014). Exploring the relationship between positive and negative emotional avoidance and anxiety symptom severity: The moderating role of attentional control. *Journal of Behavior Therapy and Experimental Psychiatry*, 45(3), 415-420.

Bartholomew, L. K., Parcel, G. S., Kok, G., Gottlieb, N. H., Fernandez, M. E. (2011) *Planning Health Promotion Programmes: An Intervention Mapping Approach.* San Francisco: Jossey-Bass.

Beck, A., Rush, A., Shaw, B., & Emery, G. (1979). *Cognitive Therapy of Depression*. New York: Guilford Press.

Biesta, G. J. J. (2010). Pragmatism and the philosophical foundations of mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), SAGE handbook of mixed methods in social & behavioral research (2nd ed., pp. 95-118). Thousand Oaks, CA: Sage.

Bohnke, J. R., & Croudace, T. J. (2016). Calibrating well-being, quality of life and common mental disorder items: psychometric epidemiology in public mental health research. *British Journal of Psychiatry*, *209*, 162-168.

Borckardt , J. J. (2006) Simulation Modelling Analysis Version 8.3.3: user's guide. Retrieved October 19, 2017, from http://www.clinicalresearcher.org/SMA_Guide.pdf.

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: A guide to case-based time-series analysis. *American Psychologist*, *63*, 77.

Brown, T. A. (2007). Temporal course and structural relationships among dimensions of temperament and DSM-IV anxiety and mood disorder constructs. *Journal of Abnormal Psychology*, *116*(2), 313.

Bryant, F. B., & Veroff, J. (2007). *Savoring: A New Model of Positive Experience*. Mahwah, NJ: Lawrence Erlbaum

Burr, L. A., Javiad, M., Jell, G., Werner-Seidler, A., & Dunn, B. D. (2017). Turning lemonade into lemons: Dampening appraisals reduce positive affect and increase negative affect during positive activity scheduling. *Behaviour Research and Therapy*, 91, 91-101.

Carver, C. S., & White, T. L. (1994) Behavioural inhibition, behavioural activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, 67, 319-333.

Chakhssi, F., Kraiss, J. T., Sommers-Spijkerman, M., Bohlmeijer, E. T. (2018). The effect of positive psychology interventions on well-being and distress in clinical samples: a systematic review and meta-analysis. *BMC Psychiatry, 18*, 211.

Chaves, C., Lopez-Gomez, I., Hervas, G., & Vazquez, C. (2017). A comparative study on the efficacy of a positive psychology intervention and a cognitive behavioural therapy for clinical depression. *Cognitive Therapy and Research, 41*, 417-433.

Clark, D.M. (2004). Developing new treatments: on the interplay between theories, experimental science and clinical innovation. *Behaviour Research and Therapy, 42*, 1089-1104.

Concannon, T. W., Meissner, P., Grunbaum, J. A., McElwee, N., Guise, J., Santa, J., Conway, P. H., Daudelin, D., Morratto, E., Leslie, L. K. (2012). A new taxonomy for stakeholder engagement in patient-centred outcomes research. *Journal of General Internal Medicine, 27*, 985-991.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). New York: Academic Press.

Cooper, R. (1999). The invisible success factors in product innovation. *Journal of Product Innovation Management, 16*, 115-133.

Corrigan, P. W., Salzer, M., Ralph, R. O., Sangster, Y., & Keck, L. (2004). Examining the factor structure of the recovery assessment scale. *Schizophrenia Bulletin*, 30, 1035-1041.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *British Medical Journal, 337*, a1655.

Craske, M., Neuret, A. E., Ritz, T., Treanor, M., Dour, H., & Rosenfield, D. (2019). Positive affect treatment for depression and anxiety. A randomized controlled trial for a core feature of anhedonia. *Journal of Consulting and Clinical Psychology, 87,* 457-471. .

Crawford, J. R., Gaithwaite, P. H., Larie, C. J., Henry, J. D., MacDonald, M. A., Sutherland, J., & Sinha, P. (2009). A convenient method of obtaining percentile norms and accompanying interval estimates for self-report mood scales (DASS, DASS-21, HADS, PANAS, and sAD). *British Journal of Clinical Psychology, 48*, 163-180.

Cuijpers, P., Karyotaki, E., Weitz, E., Andersson, G., Hollon, S. D., & van Straten, A. (2014b). The effects of psychotherapies for major depression in adults on remission, recovery and improvement: a meta-analysis. *Journal of Affective Disorders, 159*, 118-126.

Cuijpers, P., Turner, E., Mohr, D., Hofmann, S., Andersson, G., Berking, M., & Coyne, J. (2014a). Comparison of psychotherapies for adult depression to pill placebo control groups: a meta-analysis. *Psychological Medicine, 44*, 685-695.

Cuijpers, P., van Straten, A., Andersson, G., & van Oppen, P. (2008). Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. *Journal of Consulting and Clinical Psychology, 76*, 909-922.

De Shazer, S., & Coulter, M. (2012). *More than Miracles: The State of the Art of Solution-Focused Brief Therapy*. Routledge.

Demyttenaere, K., Donneau, A. F., Albert, A., Ansseau, M., Constant, E., & Van Heeringen, K. (2015). What is important in being cured from depression? Discordance between physicians and patients (1). *Journal of Affective Disorders*, *174*, 390-396.

DeRubeis, R. J., Hollon, S. D., Amsterdam, J. D., Shelton, R. C., Young, P. R., Salomon, R. M., ... & Gallop, R. (2005). Cognitive therapy vs medications in the treatment of moderate to severe depression. *Archives of General Psychiatry*, *62*, 409-416.

Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behaviour Therapy and Experimental Psychiatry, 31*, 73-86.

Dunn B. D., (In Press) Augmenting Cognitive Behavioural Therapy to build positive mood in depression. *Oxford Handbook of Positive Emotion and Psychopathology*.

Dunn, B. D. (2017). Opportunities and challenges for the emerging field of positive emotion regulation: a commentary on the special edition on positive emotions in psychopathology. *Cognitive Therapy and Research, 41,* 469-478.

Dunn, B. D. (2012). Helping depressed clients reconnect to positive emotion experience: current insights and future directions. *Clinical Psychology & Psychotherapy*, *19*, 326-340.

Dunn, B. D., & Roberts, H. (2016). Improving the capacity to treat depression using talking therapies: Setting a positive clinical psychology agenda. In A. M. Wood & J. Johnson (Eds.), *Wiley Handbook of Positive Clinical Psychology* (pp. 181-204). Chichester, UK: Wiley

Dunn, B. D., Burr, L. A., Smith, H. B., Hunt, A., Dadgostar, D., Dalglish, L., ... & Bos, N. (2018). Turning gold into lead: Dampening appraisals reduce happiness and pleasantness and increase sadness during anticipation and recall of pleasant activities in the laboratory. *Behaviour Research and Therapy*, *107*, 19-33.

Dunn, B. D., Stefanovitch, I., Buchan, K., Lawrence, A. D., & Dalgleish, T. (2009). A reduction in positive self-judgment bias is uniquely related to the anhedonic symptoms of depression. *Behaviour Research and Therapy*, *47*(5), 374-381.

Dunn, B.D., German, R. E., Kazanov, G., Zu, C., Hollon, S. D. & DeRubeis, R. J. (submitted) Changes in positive and negative affect during pharmacological treatment and cognitive therapy for major depressive disorder - A secondary analysis of two randomized controlled trials.

Dunn, B. D., Widnall, E. W., Reed, N., Taylor, R. S., Owens, C., Spencer, A…..Kuyken, W. (2019). Evaluating Augmented Depression Therapy (ADepT): Study protocol for a pilot randomised controlled trial. *Pilot and Feasibility Trials, 5*, 63.

Elliot, A. J., Sheldon, K. M., & Church, M. A. (1997). Avoidance personal goals and subjective well-being. *Personality and Social Psychology Bulletin*, *23*, 915–927.

Fava, G. A (2016). *Well-Being Therapy: treatment Manual and Clinical Applications*. Karger

Feilzer, M. Y. (2010). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research, 4*, 6-16.

Feldman, G. C., Joormann, J., & Johnson, S. L. (2008). Responses to positive affect: A self-report measure of rumination and dampening. *Cognitive Therapy and Research*, *32*(4), 507.

Fernandez, E., Salem, D., Swift, J., Ramtahahl, N. (2015). Meta-analysis of dropout from cognitive behavior therapy: Magnitude, timing and moderators. *Journal of Consulting and Clinical Psychology*, *83*, 1108-1122

First, M.B., Spitzer, M., Gibbon, M., Williams, J. B. W. (1994) *Structured Clinical Interview for DSM-IV Axis I Disorders: Clinician Version*, Washington DC: American Psychiatric Association.

Flynn, T. N., Huynh, E., Peters, T. J., Al-Janabi, H., Clemens, S., Moody, A., & Coast, J. (2015). Scoring the ICECAP-A capability instrument. Estimation of a UK general population tariff. *Health Economics*, *24*(3), 258-269.

Fournier, J. C., DeRubeis, R. J., Hollon, S. D., Dimidjian, S., Amsterdam, J. D., Shelton, R. C., & Fawcett, J. (2010). Antidepressant drug effects and depression severity: a patient-level meta-analysis. *Journal of the American Medical Association, 303*, 47-53.

Franken, I. H., Rassin, E., & Muris, P. (2007). The assessment of anhedonia in clinical and non-clinical populations: further validation of the Snaith–Hamilton Pleasure Scale (SHAPS). *Journal of Affective Disorders*, *99*(1-3), 83-89.

Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology*, 13,117-124.

Gray, J. A. (1987). *The Psychology of Fear and Stress* (2nd ed.). Cambridge: Cambridge University Press.

Geschwind, N., Arntz, A., Bannink, F., & Peeters, F. (2019). Positive cognitive behaviour therapy in the treatment of depression: A randomized order within-subject comparison with traditional cognitive behaviour therapy. *Behaviour Research and Therapy, 116*, 119-130.

Ghielen, S. T. S., van Woerkom, M., & Christina, M. Promoting positive outcomes through strengths interventions: A literature review. *The Journal of Positive Psychology, 13*, 573-585.

Gilbert, P., McEwan, K., Catarino, F., Baião, R., & Palmeira, L. (2014). Fears of happiness and compassion in relationship with depression, alexithymia, and attachment security in a depressed sample. *British Journal of Clinical Psychology*, *53*, 228-244.

Hans, E. & Hiller, W. (2013) Effectiveness and dropout from outpatient cognitive behavioural therapy for adult unipolar depression: A meta-analysis of nonrandomized effectiveness studies. *Journal of Consulting and Clinical Psychology, 81*, 75-88.

Harvey, A.G., Lee, J., Williams, J., Hollon, S., Walker, M.P., Thompson, M. & Smith, R. (2014). Improving Outcome of Psychosocial Treatments by Enhancing Memory and Learning. *Perspectives in Psychological Science, 9*, 161-179.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2011). *Acceptance and Commitment Therapy: The Process and Practice of Mindful Change*. Guilford Press.

Hofmann, S. G., Grossman, P., & Hinton, D. E. (2011). Loving-kindness and compassion meditation: Potential for psychological interventions. *Clinical Psychology Review*, 31, 1126-1132.

Hollon, S. D., DeRubeis, R. J., Fawcett, J., Amsterdam, J. D., Shelton, R. C., Zajecka, J., ... & Gallop, R. (2014). Effect of cognitive therapy with antidepressant medications vs antidepressants alone on the rate of recovery in major depressive disorder: a randomized clinical trial. *JAMA psychiatry*, *71*, 1157-1164.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., ... & Wang, P. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *American Journal of Psychiatry, 167*, 748–51.

Jacob, G. A., Ower, N., & Buchholz, A. (2013). The role of experiential avoidance, psychopathology, and borderline personality features in experiencing positive emotions: A path analysis. *Journal of Behavior Therapy and Experimental Psychiatry, 44*, 61-68.

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

Judd, L. L. (1997). The clinical course of major depressive disorders. *Archives of General Psychiatry, 54*(11), 989 - 991.

Kanter, J. W., Mulick, P. S., Busch, A. M., Berlin, K. S., & Martell, C. R. (2007). The Behavioral Activation for Depression Scale (BADS): psychometric properties and factor structure. *Journal of Psychopathology and Behavioral Assessment*, *29*(3), 191.

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings* (2nd ed.). New York, NY: Oxford University Press.

Keller, M. B., Klein, D. N., Hirschfeld, R. M., & Kocsis, J. H. (1995). Results of the DSM-IV mood disorders field trial. *The American Journal of Psychiatry*, *152*, 843-849.

Keller, M. B., Lavori, P. W., Friedman, B., Nielsen, E., Endicott, J., McDonald-Scott, P., & Andreasen, N. C. (1987). The Longitudinal Interval Follow-up Evaluation: a comprehensive method for assessing outcome in prospective longitudinal studies. *Archives of General Psychiatry, 44*(6), 540-548.

Kessler, R. C., Barber, C., Beck, A., Berglund, P., Cleary, P. D., McKenas, D., ... & Wang, P. (2003). The world health organization health and work performance questionnaire (HPQ). *Journal of Occupational and Environmental Medicine*, *45*, 156-174.

Keyes, C. L. (2005). Mental illness and/or mental health? Investigating axioms of the complete state model of health. *Journal of Consulting and Clinical Psychology*, *73*, 539.

Kitchenham, A. D. (2010). Mixed methods in case study research. In A. J. Mills, G. Durepos, & E. Wiebe (Eds.), *Encyclopedia of Case Study Research* (pp. 562-564). Thousand Oaks, CA: Sage. doi:10.4135/9781412957397.n208

Kratochwill, T. R., Levin, J. R. (2010). Enhancing the scientific credibility of single-case intervention research: randomization to the rescue. *Psychological Methods, 15*, 12-44.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-case designs technical documentation*. Washington, DC: What Works Clearinghouse, Institute of Education Sciences.

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education, 34*, 26-38.

Kring, A. M., Persons, J. B., & Thomas, C. (2007). Changes in affect during treatment for depression and anxiety. *Behaviour Research and Therapy, 45*(8), 1753-1764.

Kroenke K, Spitzer RL, Williams JB (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*, 606-613.

Kroenke, K., Spitzer, R. L., Williams, J. B. W., & Lowe, B. (2010). The patient health questionnaire somatic, anxiety and depressive symptom scales: a systematic review. *General Hospital Psychiatry, 32*, 345-359.

Kruijshaar, M. E., Barendregt, J., Vos, De Graaf, T., Spijker, R., Andrews, J. (2005) Lifetime prevalence estimates of major depression: an indirect estimation method and a quantification of recall bias. *European Journal of Epidemiology, 20*, 103–111.

Kuyken, W., Padesky, C. A., & Dudley, R. (2011). *Collaborative case conceptualization: Working effectively with clients in cognitive-behavioral therapy*. New York: Guilford Press.

Kuyken, W., Padesky, C. A., & Dudley, R. (2008). The science and practise of case conceptualisation. *Behavioural and Cognitive Psychotherapy, 36*, 757-768.

Layard, R., Clark, D., Bell, S., Knapp, M., Meacher, B., Priebe, S., Turnberg, L., & Thornicroft, G. (2006). The Depression Report: A New Deal for Depression and Anxiety Disorders. Centre for Economic Performance. London: London School of Economics. Available at: http://cep.lse.ac.uk

Lundgren, T., Luoma, J. B., Dahl, J., Strosahl, K., & Melin, L. (2012). The Bull's-Eye Values Survey: A psychometric evaluation. Cognitive and Behavioural Practice, 19, 518-526.

Macleod, A. (2017). *Prospection, well-being, and mental health*. Oxford University Press.

Maheswaran, H., Weich, S., Powell, J., & Stewart-Brown, S. (2012). Evaluating the responsiveness of the Warwick Edinburgh Mental Well-Being Scale (WEMWBS): Group and individual level analysis. *Health and Quality of Life Outcomes, 10*, 156

Manea, L., Gilbody, S., McMillan, D. (2012). Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. *Canadian Medical Association Journal, 184*, e191-196.

Martell, C. R., Dimidjian, S., & Herman-Dunn, R. (2013). *Behavioral Activation for Depression: A Clinician's Guide*. New York: Guilford Press.

Matt, G. E., Vázquez, C., & Campbell, W. K. (1992). Mood-congruent recall of affectively toned stimuli: A meta-analytic review. *Clinical Psychology Review*, *12*, 227-255.

Mazzucchelli, T. G., Kane, R. T., & Ress, C. S. (2012). Behavioural activation interventions for well-being: A meta-analysis. *Journal of Positive Psychology, 5*, 105-121.

McMakin, D. L., Olino, T. M., Porta, G., Dietz, L. J., Emslie, G., Clarke, G., ... & Shamseddeen, W. (2012). Anhedonia predicts poorer recovery among youth with selective serotonin reuptake inhibitor treatment–resistant depression. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*, 404-411.

McNeish, D. (2017) Missing data methods for arbitrary missingness with small samples. *Journal of Applied Statistics, 44*, 24-39.

Melka, S. E., Lancaster, S. L., Bryant, A. R., Rodriguez, B. F., & Weston, R. (2011). An exploratory and confirmatory factor analysis of the Affective Control Scale in an undergraduate sample. *Journal of Psychopathology and Behavioral Assessment*, *33*(4), 501-513.

Michie, S., Van Stralen, M. M., & West, R. (2011). The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation Science*, *6*(1), 42.

Moore, R. G., & Garland, A. (2003). *Cognitive Therapy for Chronic and Persistent Depression*. Chichester: John Wiley & Sons.

Morley, S., & Dowzer, C.N. (2014) Manual for the Leeds Reliable Change Indicator: Simple Excel(tm) applications for the analysis of individual patient and group data. University of Leeds, Leeds, UK.

Morley, S. (2017). Single-case methods in Clinical Psychology: A practical guide. Routledge, UK.

Moussavi, S., Chatterji, S., Verdes, E., Tandon, A., Patel, V., Ustun, B.. (2007). Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *Lancet, 370*, 851–858

Muller, M. J. (2012) Participatory Design: The third space in HCI. In "The human computer interaction handbook: Fundamentals, evolving technologies and emerging applications". New Jersey: Lawrence Erlbaum Associates.

Mundt, J. C., Marks, I. M., Shear, M. K., & Greist, J. M. (2002). The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *The British Journal of Psychiatry*, *180*, 461-464.

Nakonezny, P. A., Carmody, T. J., Morris, D. W., Kurian, B. T., & Trivedi, M. H. (2010). Psychometric evaluation of the Snaith–Hamilton Pleasure Scale (SHAPS) in adult outpatients with major depressive disorder. *International Clinical Psychopharmacology*, *25*, 328.

Naragon-Gainey, K., Gallagher, M. W., & Brown, T. A. (2013). Stable "trait" variance of temperament as a predictor of the temporal course of depression and social phobia. *Journal of Abnormal Psychology*, *122*, 611.

O'Cathain, A., Thomas, K. J., Drabble, S. J., Rudolph, A., & Hewison, J. (2013) What can qualitative research do for randomised controlled trials? A systematic mapping review. *BMJ Open, 3*,e002889.

Onghena, M., Maes, B., Heyvaert, M.(2018). Mixed Methods Single Case Research: State of the Art and Future Directions. *Journal of Mixed Methods Research*, https://doi.org/10.1177/1558689818789530

Padesky, C. A., & Mooney, K. A. (2012). Strengths-based cognitive–behavioural therapy: A four-step model to build resilience. *Clinical Psychology & Psychotherapy*, *19*(4), 283-290.

Paulus, M. P., Stein, M. B., Craske, M. G., Bookheimer, S., Taylor, C. T., Simmons, A. N……, Fan, B.,(2017). Latent variable analysis of positive and negative valence processing focused on symptom and behavioural units of analysis in mood and anxiety disorders. *Journal of Affective Disorders, 216*, 17-29.

Pelizza, L., & Ferrari, A. (2009). Anhedonia in schizophrenia and major depression: state or trait? *Annals of General Psychiatry, 8*, 1-9.

Plummer, F., Manea, I., Trepel, D., & McMillan, D. (2016). Screening for anxiety disorders with the GAD-7 and GAD-2: A systematic review and diagnostic metaanalysis. General Hospital Psychiatry, 39, 24-31.

Powell, J., Hamborg, T., Stallard, N., Burls, A., McSorley, J., Bennett, K., Griffiths, K. M., & Christensen, H. (2013). Effectiveness of a web-based cognitive-behavioural tool to improve mental well-being in the general population: Randomized Controlled Trial. *Journal of Medical Internet Research, 15*: e2.

Richards, D. A., Ekers, D., McMillan, D., Taylor, R. S., Byford, S., Warren, F. C., ... & O'Mahen, H. (2016). Cost and Outcome of Behavioural Activation versus Cognitive Behavioural Therapy for Depression (COBRA): a randomised, controlled, non-inferiority trial. *The Lancet*, *388*, 871-880.

Richie, J., Spencer, L. and O'Connor, W. (2003) *Qualitative Research Practice*. London: Sage Publication.

Roberts, B. W., Luo, J., Briley, D. A., Chow, P. I., Su, R., & Hill, P. L. (2017). A systematic review of personality change through intervention. *Psychological Bulletin, 143*, 117-141.

Roskes, M., Elliot, A. J., & De Dreu, C. K. W. (2014). Why is avoidance motivation problematic, and what can be done about it? *Current Directions in Psychological Science, 23*, 133-138.

Ruini, C., & Fava, G. A. (2012). Role of well-being therapy in achieving a balanced and individualized path to optimal functioning. *Clinical Psychology & Psychotherapy*, *19*, 291-304.

Rush, A. J. (2015). Distinguishing functional from syndromal recovery: Implications for clinical care and research. *Journal of Clinical Psychiatry, 76*, e832-e834.

Segal, Z. V., Williams, M., & Teasdale, J. D. (2018). *Mindfulness-based cognitive therapy for depression*. New York: Guilford Press.

Sheehan, D. V., Nakagome, K., Asami, Y., Pappadopulos, E. A., & Boucher, M. (2017). Restoring function in major depressive disorder: A systematic review. *Journal of Affective Disorders, 215*, 299-313.

Sheehan, D. V., Harnett-Sheehan, K., Spann, M. E., Thompson, H. F., & Prakash, A. (2011). Assessing remission in major depressive disorder and generalized anxiety disorder clinical trials with the discan metric of the Sheehan disability scale. *International Clinical Psychopharmacology, 26*, 75-83.

Sin, N. L., & Lyubomirsky, S. (2009). Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: A practice-friendly meta-analysis. *Journal of Clinical Psychology, 65*(5), 467-487.

Singleton, N., Bumpstead, R., O'Brien, M., Lee, A., & Meltzer, H. (2003). Psychiatric morbidity among adults living in private households, 2000. *International Review of Psychiatry, 15*(1-2), 65-73.

Slade, M. (2010). Mental illness and well-being: the central importance of positive psychology and recovery approaches. *BMC Health Services Research, 10*, 26.

Smith, J. D. (2012). Single-case experimental designs: A systematic review of published research and current standards. *Psychological Methods, 17*, 510-550.

Snaith, R. P., Hamilton, M., Morley, S., Humayan, A., Hargreaves, D., & Trigwell, P. (1995). A scale for the assessment of hedonic tone the Snaith–Hamilton Pleasure Scale. *The British Journal of Psychiatry, 167*, 99-103.

Spijker, J., Bijl, R. V., De Graaf, R., & Nolen, W. A. (2001). Determinants of poor 1-year outcome of DSM-III-R major depression in the general population: results of the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Acta Psychiatrica Scandinavica, 103*, 122-130.

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine, 166*, 1092-1097.

Stewart-Brown, S. L., Platt, S., Tennant, A., Maheswaran, H., Parkinson, J., Weich, S., ... & Clarke, A. (2011). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): a valid and reliable tool for measuring mental well-being in diverse populations and projects. *Jounral of Epidemiology and Community Health*, *65*(Suppl 2), A38-A39.

Stewart-Brown, S., Tennant, A., Tennant, R., Platt, S., Parkinson, J., & Weich, S. (2009). Internal construct validity of the Warwick-Edinburgh mental well-being scale (WEMWBS): a Rasch analysis using data from the Scottish health education population survey. *Health and Quality of Life Outcomes*, *7*:15

Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Barlow, D…… Wilson, B. (2016). The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016 statement. *Archives of Scientific Psychology, 4*, 1-9.

Taylor, C. T., Lyubomirsky, S., & Stein, M. B. (2017). Upregulating the positive affect system in anxiety and depression: Outcomes of a positive activity intervention. *Depression & Anxiety, 34*, 267-280.

Teasdale, J. D. (1999). Emotional processing, three modes of mind and the prevention of relapse in depression. *Behaviour Research and Therapy, 37*, 53–77.

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., ... & Stewart-Brown, S. (2007). The Warwick-Edinburgh mental well-being scale (WEMWBS): development and UK validation. *Health and Quality of life Outcomes, 5*, 63.

Trompetter, H. R., Lamers, S. M. A., Westerhof, G. J., Fledderus, M., & Bohlmeijer, E. T. (2017). Both positive mental health and psychopathology should be monitored in psychotherapy: confirmation for the dual-factor model in acceptance and commitment therapy. *Behaviour Research and Therapy*, *91*, 58-63.

Üstün, T., Ayuso-Mateos, J. L., Chatterji, S., Mathers, C., & Murray, C. J. (2004). Global burden of depressive disorders in the year 2000. *The British Journal of Psychiatry, 184*, 386-392.

Verplanken, B. (2006). Beyond frequency: Habit as mental construct. *British Journal of Social Psychology*, *45*, 639-656.

Vilhauer, J. (2014). *Think forward to thrive: How to use the mind's power of anticipation to transcend your past and transform your life*. Calafornia: New World Library.

Vittengl, J. R., Clark, L. A., Dunn, T. W., & Jarrett, R. B. (2007). Reducing relapse and recurrence in unipolar depression: a comparative meta-analysis of cognitive-behavioral therapy's effects. *Journal of Consulting and Clinical Psychology, 75*, 475-488.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063-1070.

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology*, *76*, 820-838..

Watson, D., & Clark, L. A. (1999). The PANAS-X: Manual for the Positive and Negative Affect Schedule – Expanded Form. University of Iowa (unpublished)

Werner-Seidler, A., Banks, R., Dunn, B. D., & Moulds, M. L. (2013). An investigation of the relationship between positive affect regulation and depression. *Behaviour Research and Therapy*, *51*, 46-56.

Widnall, E., Price, A., Trompetter, H., & Dunn, B. D. (submitted). Routine Cognitive Behavioural Therapy for anxiety and depression is more effective at repairing symptoms of psychopathology than enhancing wellbeing.

Wiles, N., Thomas, L., Ridgway, N., Turner, N., Campbell, J., Garland, A.,…..& Lewis, G. (2013). Cognitive behavioural therapy as an adjunct to pharmacotherapy for primary care based patients with treatment resistant depression: results of the COBALT randomised controlled trial. *Lancet, 381*, 375-384.

Williams, K. E., Chambless, D. L., & Ahrens, A.H. (1997). Are emotions frightening? An extension of the fear of fear concept. *Behaviour Research and Therapy,* 35, 239-248.

Winer, E. S., & Salem, T. (2016). Reward devaluation: Dot-probe meta-analytic evidence of avoidance of positive information in depressed persons. *Psychological Bulletin, 142,* 18-78.

Wood, A.M., & Tarrier, N. (2010) Positive Clinical Psychology: A new vision and strategy for integrated research and practice. *Clinical Psychology Review, 30,* 819–829.

Wood, W., & Neal, D. T. (2007). A new look at habits and the habit-goal interface. *Psychological Review, 114*, 843-863.

Zahra, D., Qureshi, A., Henley, W., Taylor, R., Quinn, C., Pooler, J., ... & Byng, R. (2014). The work and social adjustment scale: reliability, sensitivity and value. *International Journal of Psychiatry in Clinical Practice, 18*, 131-138.

Zimmerman, M., McGlinchey, J. B., Posternak, M. A., Friedman, M., Attiullah, N., & Boerescu, D. (2006). How should remission from depression be defined? The depressed patient's perspective. *American Journal of Psychiatry, 163*, 148-150.

Table 1
Participant demographic and clinical characteristics

P.	Age	Sex	PHQ	GAD-7	WEM-WBS	WSAS	SHAPS	PA	NA	First Onset	# Episodes	Therapist	Previous therapy	Medication Status	Employment
3	47	F	23	16	29	25	37	15	42	13	10*	1	C, MBCT	Duloxetine 60mg, diazepam 2 mg[a]	U
4	34	M	14	16	29	28	29	14	28	18	7*	2	C, CBT	Sertraline 50mg, propanalol 10mg	U
5	27	F	26	20	36	33	28	27	45	12	10	3	LI- CBT	Citalopram 20mg, propanalol 40mg[b]	U
6	43	M	11	8	34	28	34	24	28	13	8*	2	LI-CBT	Sertraline 150mg	U
7	58	F	13	9	37	31	22	20	23	15	15	3	None	Duloxetine 60mg[c]	R
8	48	F	20	14	34	28	34	23	31	46	3	3	C, LI-CBT	None	E
9	33	M	11	7	28	26	30	18	23	27	2	1	LI-BA	Fluoxetine 20mg	E
10	31	F	19	12	25	34	44	11	36	18	10	2	LI-CBT	Fluoxetine 20 mg	E
11	64	F	17	16	32	-	36	12	25	50	10	3+	C	None	R
12	22	F	13	17	33	23	31	20	33	12	10	1	LI-BA	Fluoxetine 20mg	E
13	21	M	19	13	23	32	32	16	33	21	2	2	LI-BA	None	S
M	38.91	-	16.91	13.45	30.90	28.80	32.45	18.18	31.54	22.27	7.91	-	-	-	-
SD	14.24	-	4.97	4.11	4.48	3.62	5.66	5.13	7.27	13.52	4.09	-	-	-	-

Note- P. = participant id; CBT; Previous therapy: C = counselling; MBCT = Mindfulness Based Cognitive Therapy; LI-CBT = low intensity CBT; LI-BA = low intensity BA. Therapist: + booster sessions completed by therapist 1, due to therapist 3 becoming ill. Medication status: a - Duloxetine replaced with mirtazapine 15mg mid treatment; [b] - Stopped taking citalopram but continued with propanalol mid treatment; [c]- Reduced dose of duloxetine to 40mg mid treatment; * = previous suicide attempt. Employment: U = unemployed; R = retired; E = employed; S= student.

Table 2

Sessions attended and client Ratings of treatment acceptability, satisfaction with treatment, and whether they would recommend treatment to friends and family

Measure	3	4	5	6	7	8	9	10	11	12	13	Mean (SD)
Acute sessions attended (out of maximum 15)	15	15	15	15	15	15	11	15	15	15	15	14.64 (1.21)
Booster sessions attended (out of maximum 5)	2	0	5	1	1	5	0	4	1	0	3	2.00 (1.94)
Acceptability ratings	4	-	5	5	4	5	4	4	5	4	-	4.44 (0.53)
Satisfaction ratings	4	-	5	4	4	5	4	4	5	5	-	4.44 (0.53)
Recommend ratings	5	-	5	5	5	5	5	4	5	5	-	4.89 (0.33)

Note –Acceptability of the intervention rated from 1 (not at all acceptable) to 5 (extremely acceptable); satisfaction with intervention rated from 1 (not at all satisfied) to 5 (extremely satisfied); and whether they would recommend this treatment to friends or family suffering from depression rated from 1 (not at all likely) to 5 (very likely).

Table 3

Time Series Analyses for Patient Health Questionnaire (PHQ-9) and the Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS)

	3	4	5	6	7	8	9	10	11	12	13
PHQ-9											
pAR(Lag1)	0.51	0.13	0.26	0.20	0.33	0.65	0.53	0.05	0.71	-0.07	0.34
Level	-.21	-.31	-.18	-.44	-.29	-.35	-.64	-.49	-.55	-.33	-.42
Slope	-.28,	-.62*	-.20	-.58*	-.58*	-.84*	-.83*	-.39	-.88*	-.19	-.70*
WEMWBS											
pAR(Lag1)	0.54	0.08	0.37	0.87	0.29	0.72	0.29	0.30	0.47	0.48	0.218
Level	.18	.29	-.01	.74	.45	.57	.52	.25	.49	.53	.27
Slope	.61	.56*	-.14	.93*	.62*	.90*	.66*	.26	.83*	.45	.58*

Note- pAR(Lag1) = autocorrelation estimate. * = fit of model significantly greater than chance at $p<.05$.

Table 4

Depression, wellbeing, anhedonia, anxiety, and functioning measures at each phase (and percentage showing reliable and clinically significant change)

	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	RC (CSC)
PHQ-9												
Pre	18	18	24	11	14	19	11	15	17	12	20	-
Post	14	8**	25	5**	9**	9**	0**	9**	0**	8	10*	73% (64%)
2mFU	16	-	26	4**	10	7**	4**	12	0**	7**	14*	60% (50%)
1yFU	10*	-	23	7	16	5**	-	9**	2**	10	-	50% (38%)
WEMWBS												
Pre	32	29	36	28	39	31	31	31	36	42	21	-
Post	36	31	31	52**	51**	45**	49**	36	63**	50**	37*	64% (55%)
2mFU	30	-	25+	53**	47**	47**	38*	32	63**	52**	38*	70% (50%)
1yFU	47**	-	34	51**	52**	55**	-	36	63**	46	-	63% (63%)
WSAS												
Pre	28	23	30	27	27	29	25	36	16	19	27	-
Post	21	11**	33	13**	23	12**	13**	24*	0**	10**	15**	73% (64%)
FU	28	-	32	6**	29	11**	13**	22*	0**	17	25	50% (40%)
1yrFU	17*	-	31	12**	31	7**	-	21*	0**	11**	-	75% (50%)
GAD-7												
Pre	15	17	19	8	9	14	7	9	15	19	12	-
Post	12	13	18	3**	4**	10	2**	7	0**	8*	8	45% (36%)
2mFU	14	-	20	2**	9	5**	4	7	0**	17	10	30% (30%)
1yrFU	9*	-	18	2**	4**	4**	-	7	0**	14*	-	75% (50%)
SHAPS												
Pre	29	27	27	32	23	39	31	43	31	26	32	-
Post	28	-	32	28	15**	20**	24	36	14**	17**	28	40% (40%)
2mFU	32	-	31	28	24	20**	26	29*	14**	19	28	30% (20%)
1yFU	24	-	27	22**	16	19**		28*	14**	20	-	50% (38%)
PA												
Pre	20	17	27	23	21	16	21	17	16	21	14	-
Post	19	-	26	30**	30**	31**	33**	20	42**	36**	24*	70% (60%)
2mFU	16	-	22+	29**	19	32**	28**	14	41**	34**	22*	60% (50%)
1yrFU	37**	-	19+	34**	26**	39**	-	23*	39**	34**	-	88% (75%)

NA												
Pre	41	28	39	25	21	41	17	32	31	36	23	-
Post	42	-	41	16**	28	20**	16**	22**	13**	16**	20**	70% (70%)
2m FU	34*	-	46	12**	21	21**	19	22**	12**	23**	22**	70% (60%)
1yrFU	28*	-	35*	18**	32+	16**	-	22**	10**	32	-	75% (50%)

Note – PHQ-9 = Patient Health Questionnaire; WEMWBS = Warwick-Edinburgh Mental Wellbeing Scale; WSAS = Work and Social Adjustment Scale; GAD-7 = Generalized Anxiety Disorder Assessment; SHAPS = Snaith Hamilton Pleasure Scale; PA = Positive affect subscale of the Positive and Negative Affect Scale; NA = Negative affect subscale of the Positive and Negative Affect Scale. * = reliable change (relative to pre-assessment); ** = reliable and clinically significant change (relative to pre-assessment); + = reliable deterioration. RC (CSC) = % of participants with data showing reliable change (clinically significant change). Clinically significant change used criterion c for PHQ-9, WEMWBS, GAD-7, SHAPS, PA and NA and criterion a for WSAS (as comparison sample normative data are not available for this measure).

Figure 1
Logic Model of ADepT

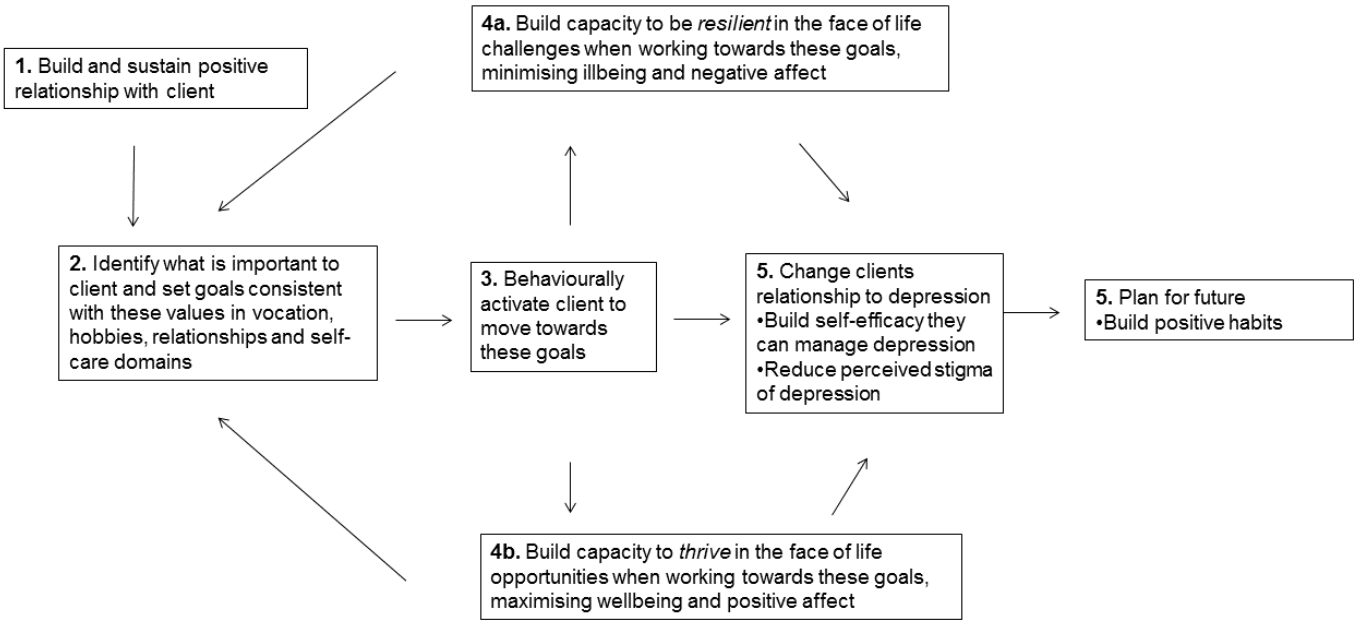


Figure 2
Consort Diagram for Case Series

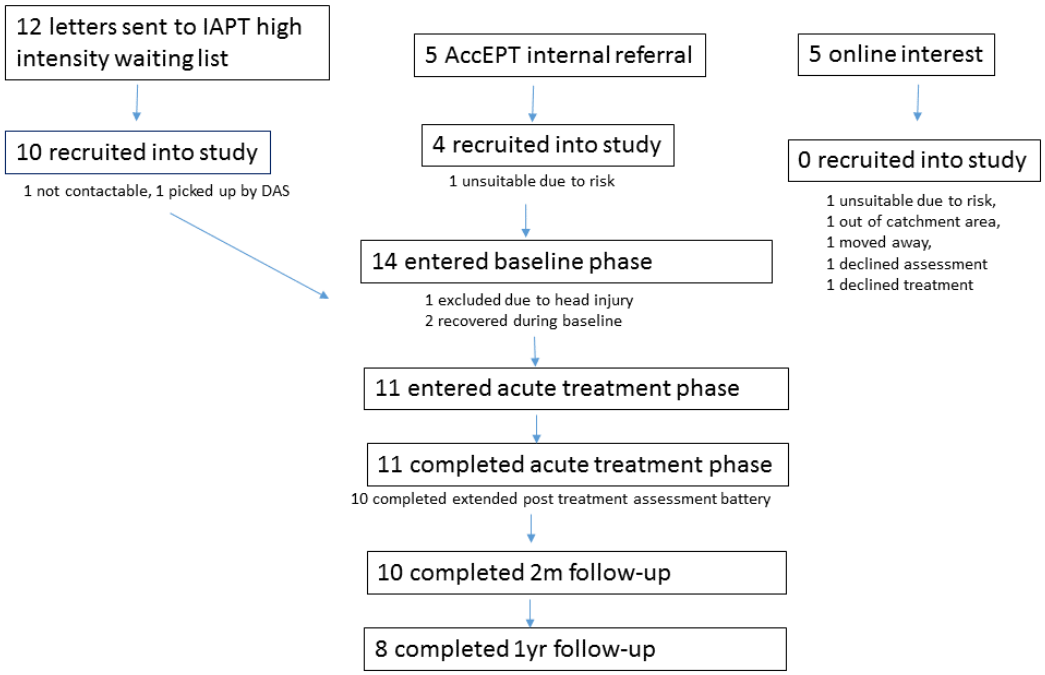
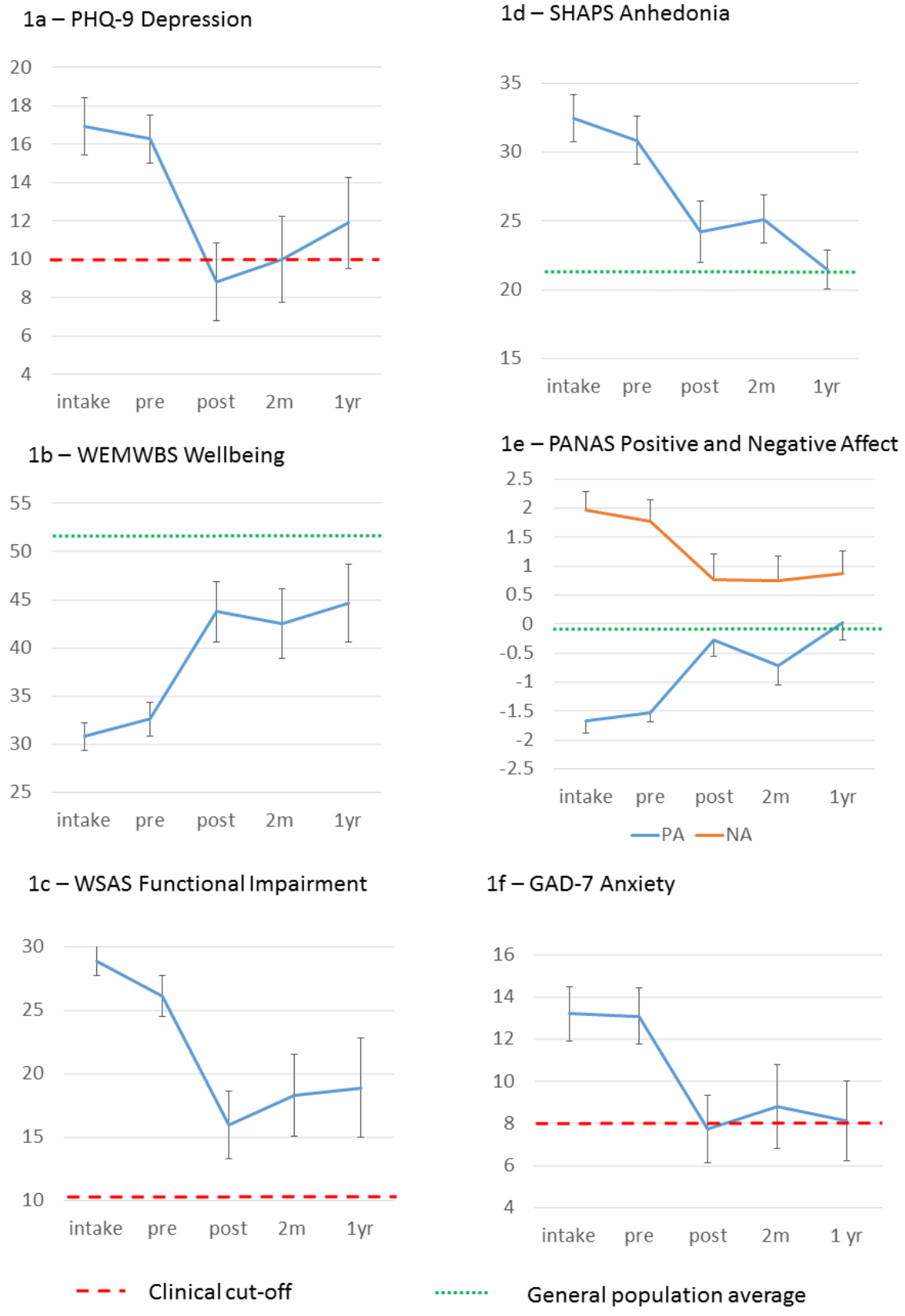


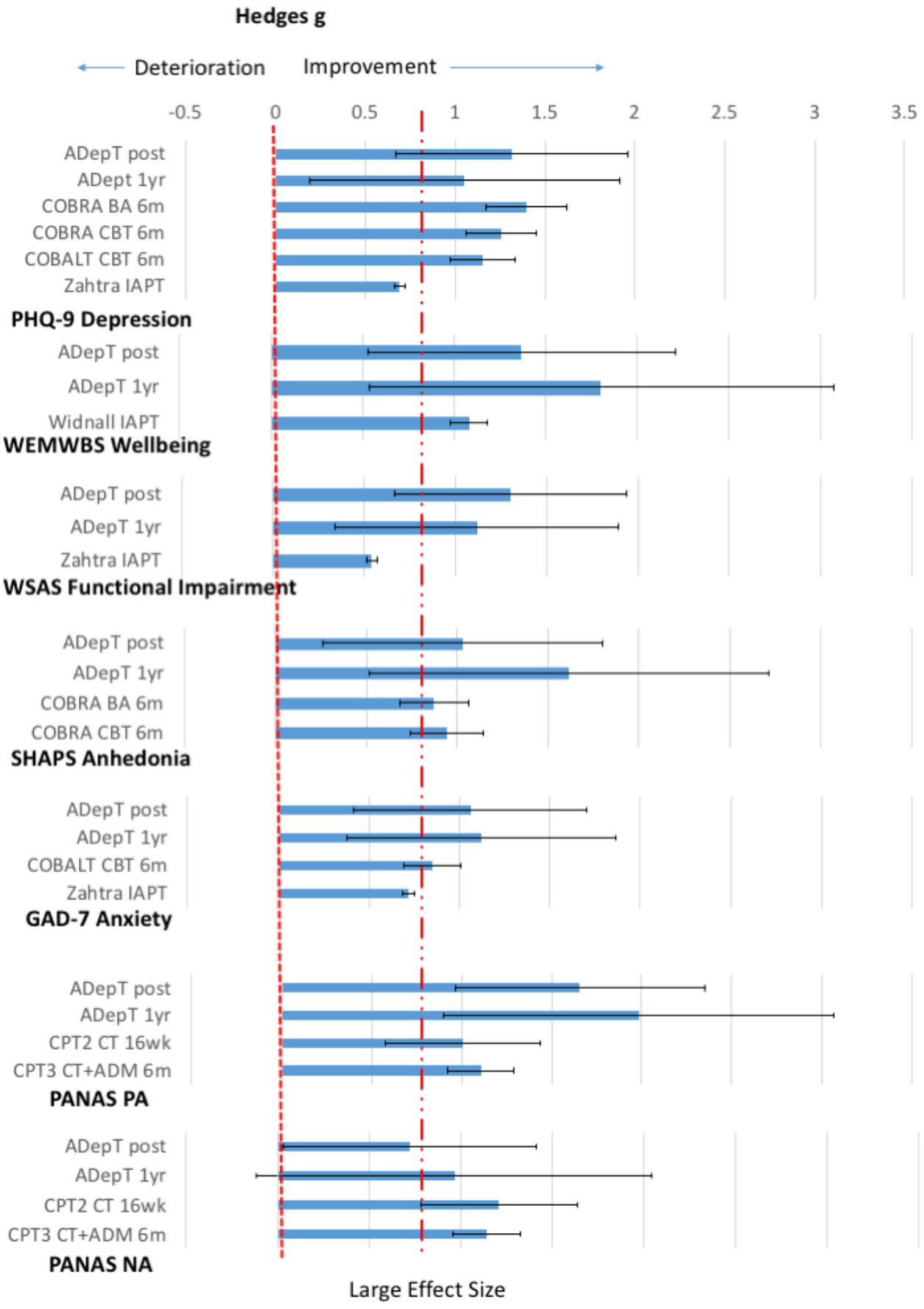
Figure 3

Mean scores for depression (a), wellbeing (b), WSAS (c), SHAPS anhedonia (d), PANAS positive and negative affect (e) and anxiety (f) at each assessment point



Note – Data are mean (one SEM values). Y axes are raw score values, with the exception of the PANAS which is Z score units (to allow visual comparison of repair of NA versus PA). Raw plots of the PANAS data are Supplementary Figure 2.

Figure 4
Benchmarking ADepT effect sizes against other datasets



Note – Effect size estimates are Hedges g_{av} (using small sample correction), with 95% confidence intervals.

Footnotes

1. Similar findings emerge when looking at the broader constructs of wellbeing versus symptom relief and personality change. For example, there is more robust relief of symptoms of mood disorders than improvements in wellbeing following Acceptance and Commitment Therapy (Trompetter et al., 2017) and CBT (Widnall et al., submitted). Behavioural Activation only generates a medium effect size for wellbeing outcomes (Mazzuchelli et al, 2012). Moreover, psychological treatments bring about bigger changes in neuroticism (related to negative affect) than they do extraversion (related to positive affect) (see systematic review by Roberts et al., 2017).
2. This extended battery also included the Behavioural Activation for Depression Scale (BADS) as a measure of behavioural activation, the short-form of the Recovery Assessment Scale (RAS-SF) as a measure of recovery orientation, the response to positive affect scale (RPA) as measure of positive appraisal, the positive scale of the Affective Control Scale (ACS-P) to measure fear of positive affect, and the ICEPOP capability measure for adult (ICECAP-A) as an additional measure of wellbeing. For the sake of brevity, analyses of these data are reported in the supplementary materials (Section S3). There were significant improvements during treatment on the BADS, RAS-SF, ICECAP-A, and dampening scale of the RPA, which were largely sustained over one-year follow-up. There was no significant change in fear of positive emotions on the ACS-P, or amplifying appraisals on the RPA.
3. These three participants all continued with acute ADepT therapy but were excluded from subsequent analyses. The two participants (P1 and P2) who were excluded from the case series analyses on the basis of recovering during the baseline both completed a full dose of acute treatment and rated it is as acceptable. They were both in recovery for depression and anxiety at the end of treatment. The participant with a head injury (P14) completed a modified treatment, before being referred onto specialist head injury services.

S1 Full Description of Outcome Measures Used

The Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS, Tennant et al., 2007) was used to measure wellbeing. Participants rate to what extent they have felt the way described by 14 statements about wellbeing (e.g. "I've been feeling optimistic about the future"), on a scale ranging from 1 (none of the time) to 5 (all of the time). Scores range from 14 (low wellbeing) to 70 (high wellbeing). The WEMWBS has been shown to have good internal reliability (Cronbach's $\alpha = 0.91$) and test-retest reliability (intraclass correlation = 0.83) (Tennant et al., 2007; Stewart-Brown et al., 2009; 2011). In 7020 adults from the UK general population the mean score was 51.61 (SD=8.71) (taken as part of the Health Survey for England, 2011; see https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/using/howto/)

The Patient Health Questionnaire (PHQ-9; Kroenke et al., 2001) was used to measure depression symptom severity. Participants rate the frequency over the past two weeks with which they have experienced the nine DSM-V symptoms of depression (e.g. "little interest or pleasure in doing things), on a scale ranging from 0 (not at all) to 3 (nearly every day). Scores range from 0 (asymptomatic) to 27 (severely depressed). A cut of score of $\geq$ 10 has been found to be a good proxy for meeting diagnostic criteria for a major depressive episode as measured by structured clinical interview, with 88% sensitivity and 88% specificity (Kroenke et al., 2010). Studies find the PHQ-9 has good internal reliability ($\alpha > .86$) and test-retest reliability (intraclass correlation=.84) (Kroenke et al., 2001). In 5018 adults from the German general population, the mean score was 2.91 (SD=3.52)(Kocalevent, Hinz & Brahler, 2013).

The Work and Social Adjustment Scale (WSAS; Mundt et al., 2002) was used to assess functional impairment resulting from depression. Participants rate to what extent their problem (in this case depression) gets in the way of their ability to work, to manage the home, engage with social leisure activities, engage with private leisure activities, and form and maintain close relationships. Each of five items is rated on a scale from 0 (not at all) to 8 (very severely), leading to a total score ranging from 0 (no impairment) to 40 (extreme impairment). Scores <10 are associated with subclinical populations, scores between 10 and 19 suggest significant functional impairment but less severe clinical symptomatology, and scores above 20 suggest moderately severe or worse psychopathology. The WSAS has adequate internal

(α=0.84; Zahra et al., 2014) and test-retest (two week intraclass correlation=0.73; Mundt et al., 2002) reliability. General population data are not available for the WSAS.

The past week version of Positive and Negative Affect Scale (PANAS; Watson et al., 1988) was used to measure levels of positive and negative affect (measures of PVS and NVS respectively). Participants rate to what extent 20 adjectives (ten positive, e.g. 'excited', 'interested'; ten negative, e.g. 'afraid', 'scared') describe how they have felt on average over the past week, on a scale ranging from 1 (Very slightly or not at all) to 5 (Extremely). Total scores for positive and negative affect range from 10 to 50. The PANAS factors have been found to have adequate internal reliability in past week form in adult community samples (α = .90) (Watson & Clark, 1999). Test-retest reliability is not available for the past week version of the PANAS. In 328 adults from the US general population, the mean score for the past week variant of the PANAS was 31.1 (SD=7.5) for PA and 18.0 (SD=7.1) for NA (Watson & Clark, 1999).

The Snaith Hamilton Pleasure Scale (SHAPS; Snaith et al., 1995) was to measure anhedonia severity. Participants rate over the past few days how much they agree with various statements describing their capacity to experience pleasure (e.g. 'I would enjoy a cup of tea or coffee or my favourite drink), on a scale ranging from strongly agree (1) to strongly disagree (4). We used the revised continuous scoring recommended by Franken et al (2007), where scores for each item are summed to generate a score from 14 (not at all anhedonic) to 46 (extremely anhedonic). The SHAPS has been found to have good internal reliability (α=.91) and test-retest reliability (intraclass correlation=.70) (Nakonezny et al., 2010; Franken et al. 2007). In 50 adults from the Dutch general population, the mean (revised continuous) score was 20.2 (SD=4.4) (Franken et al., 2007).

The Generalized Anxiety Disorder scale (GAD-7; Spitzer et al., 2006) was used to measure anxiety symptom severity, given depression is frequently comorbid with anxiety. Participants rate the frequency over the past two weeks with which they have experienced seven symptoms of anxiety (e.g. "feeling nervous, anxious, or on edge"), on a scale ranging from 0 (not at all) to 3 (nearly every day). Scores range from 0 (asymptomatic) to 21 (severely anxious). A cut off score of $\geq$8 has been shown to have adequate sensitivity (83%) and specificity (84%) to confirm a diagnosis of generalized anxiety disorder based on structured clinical interview (Plummer, Manea, Trepel & McMillan, 2016). The GAD-7 has been found to

have good internal reliability (α=.92) and test-retest reliability (intraclass correlation=.83) (Spitzer et al., 2006). In 5030 adults from the German general population, the mean score (averaging across men and women) was 2.93 (SD=3.38) (Lowe et al., 2008).

S2 Visual analysis of depression and wellbeing time series data

Supplementary Figures 1a to 1c plot weekly levels of depression and wellbeing during the baseline and treatment phase for each participant. Visual inspection revealed improvement in depression and wellbeing during acute treatment for seven participants (P4, P6, P7, P8, P9, P11, P13; 64%). P3 showed worsening depression and wellbeing mid-treatment, but with a slight improvement at the end of therapy. P10 showed an initial improvement in depression and wellbeing, but with a slight deterioration in the final three sessions. P5 showed no evidence of improvement during treatment for either depression or wellbeing. P12 showed no clear pattern of wellbeing change, but an improvement in depression at the end of therapy. These conclusions align with the findings from simulation-modelling analysis (SMA) presented in the main paper. However, it is important to acknowledge that there was rarely a clearly linear pattern of improvement, with many participants exhibiting depression 'spikes' and wellbeing 'drops' mid-treatment, usually linked to life events the individual was encountering during treatment. One way to interpret this is that treatment activated clients to take opportunities and face challenges, and often asked them to experiment with new ways of coping while doing so. Therefore, it makes sense for there to be considerable mood variability in the middle course of treatment. Such mood variability as a result of such exposure has been seen as a necessary part of successful recovery during depression therapy (e.g. Hayes, 2015).

S3 Analysis of Other Outcome Measures in Extended Battery

Extended Assessment Measures

The short version of the Recovery Assessment Scale (RAS-SF; Corrigan et al. 2004) was used to measure a recovery orientation. Twenty positively keyed items index a recovery orientation, each rated on a scale from 1 (strongly disagree) to 5 (strongly agree). Five subscales are generated: personal confidence and hope (e.g., 'fear doesn't stop me from living the way I want to', 9 items; scores range 9 to 45, α=.87);

willingness to ask for help (e.g., 'I know when to ask for help', 3 items, scores range 3 to 15, α=.84); goal and success orientation (e.g., 'I have a desire to succeed', 5 items, scores range 5 to 25, α=.82); reliance on others (e.g., 'Even when I don't care about myself, other people do', 4 items, scores range 4 to 20, α=.74); and relationship to symptoms (e.g., 'coping with mental illness is no longer the main focus of my life', 3 items; scores range 3 to 15, α=.74). Due to administrator error, four items were missing from the personal confidence and hope scale. We computed the scale score for the remaining 5 items, and then factored this up to align with the original scale score. We focus here on total scale score, summing together scores for each of the other factors. Higher scores indicate a greater recovery orientation. In a sample of 500 individuals suffering from mental illness, mean total RAS-SF at intake score was 87.8 (SD=14.05) (Cook et al., 2012). In 217 individuals in this sample assigned to wellness recovery action planning (an eight-week peer-led recovery initiative to promote illness self-management), mean RAS-F post-treatment was 93.7 (SD=14.7; an improvement of 4.9 points). Total scale score showed high internal reliability in this sample (α =.91; Cook et al., 2012). Test-retest reliability data are not available.

The Behavioural Activation Scale for Depression (BADS; Kanter et al., 2007) was used to index an increase in activation and decreases in avoidance, rumination and associated functional impairment. Participants are asked to rate to what extent 25 statements are true of them over the past week, on a 7 point scale ranging from 0 (not at all) to 6 (completely). BADS items measures four subscales: activation (BADS-A, 7 items; e.g., 'I did something that was hard but worth it'; α=.85), avoidance/rumination (BADS-AR, 8 items, e.g., 'I did things to avoid feeling sadness or other painful emotions', α=.86); work/school impairment (BADS-WS, e.g., 'I went to work/school, but spent my time doing other activities than my assigned tasks', 5 items, α=.76); and social impairment (BADS-S, e.g., 'I was not social, even though I had opportunities to be'; α=.82). For the activation scale, higher scores equate to better functioning. For all of the other scales, higher scores equate to worse functioning. Total scale score is computed by adding each factor, having first reversed the activation factor. The total scale score has adequate internal (α=.87) and test-retest reliability (one week intraclass correlation=.74). We focus here on the scale total score, with higher scores indicating greater Behavioural Activation. Normative data are available on 391 undergraduate

students (Kanter et al., 2007), who had a total scale mean score of 122.07 (SD=20.65). For a sample of individuals scoring above the cut-off for depression in a self-report scale, mean total score was 69.83 (SD=20.15).

The Response to Positive Affect Scale (RPA; Feldman et al. 2008) was used to measure appraisals in response to positive emotions. Participants are asked to rate 17 statements on a 4-point scale ranging from 1 (almost never) to 4 (almost always) thinking about when they feel happy, excited or enthused. The scale consists of three subscales: Emotion-focus (EF) appraisals (5 items; e.g., 'When you are feeling happy, how often do you notice you feel full of energy', $\alpha = 0.71$); dampening appraisals (8 items; e.g., 'When you are feeling happy, how often do you think "my streak of luck is going to end soon", $\alpha = 0.79$); and self-focus (SF) appraisals (4 items; e.g., 'When you are feeling happy, how often do you think "I'm achieving everything", $\alpha = 0.69$). Test-retest reliability data are not available for this scale. In a community sample of 528 respondents (Raes et al., 2009) used to validate the RPA, mean dampening score was 11.27 (SD = 3.33), mean EF score was 13.52 (SD = 2.38), and mean SF score was 9.33 (SD = 2.56).

The Positive Emotion Factor of the Affective Control Scale (ACS; Williams, Chambless & Ahrens, 1997) was used to measure fear of experiencing positive emotions. We used the revised scoring for the ACS positive scale proposed by Melka et al. (2012), which excludes the four reverse keyed items and the forward key item 10 ('Having an orgasm is scary for me because I am afraid of losing control'). Total scores on the revised 8-item scale range from 8 to 56, with higher score indicating a greater fear of experiencing positive emotions. The mean score in 1560 undergraduates (averaging across male and female participants) was 24.62 (SD = 8.37). Normative data for a depressed sample are not available for this measure. The scale has adequate internal reliability ($\alpha = .72$). Test-retest reliability data are not available.

The ICEpop CAPability Measure for Adults (ICECAP-A, Al-Janabi, Flynn & Coast, 2012) measures individuals' capability to 'do' and 'be' the things that are important in life. It has been designed to be a measure of quality of life sensitive to wellbeing for use in health economic evaluations. We included it in this study to examine if it is sensitive to wellbeing deficits in depression and change in wellbeing during treatment. For five key areas of life participants are asked to select one of four levels of capability, ranging from 1 (absence of capability) to 4 (full capability) The state 44444 would indicate full capability on all

attributes, whilst the state 11111 would indicate an absence of capability. The five over-arching attributes of capability (each example rated as four) are: stability (e.g., 'I am able to feel settled and secure in all areas of my life'), attachment (e.g., 'I can have a lot of love, friendship and support') achievement ('I can achieve and progress in all aspects of my life'), autonomy (e.g. 'I am able to be completely independent') and enjoyment (e.g. 'I can have a lot of enjoyment and pleasure'). A heterogeneity-adjusted population level tariff value (based on a UK population) for an overall state has been developed (Flynn et al., 2015). Two-week test-retest reliability of the ICECAP-A is adequate (intraclass correlation = 0.72; Al-Janabi et al., 2015). The ICECAP-A has been validated for use in depression. The average ICECAP-A score in a depressed sample (n=617 individuals) has been estimated as 0.64 (95% CI=0.62 to 0.65) and is considerably lower than the healthy general population (e.g. mean = 0.88, SD = .12 in 456 UK adults; Keeley et al., 2016).

These measures were either administered via post or in a face-to-face interview conducted by a member of the research team (EW). Table S1 reports individuals' and group average scores on each of these measures at each assessment point.

Extended Battery Scores at Intake

Intake RPA-EF and RPA-SF scores were well below, and RPA-dampening and ACS-P scores were well above, comparison sample normative data (cf. Feldman et al., 2008; Melka et al., 2012). BADS total score was also impaired relative to comparison samples and showed greater impairment than other depressed samples (M=69.83, SD=20.15 in 193 individuals scoring above caseness cut-offs for depression; Kanter et al., 2007). Intake ICECAP tariff scores were UK below general population averages (approximately .83) and also lower than estimates from other depressed samples (.64) (see Mitchell et al., 2017). This suggests the present sample at intake had significant deficits on most measures.

Reliable and clinically significant change analyses

As in the main manuscript, we looked at the number and proportion of individuals showing reliable change (RC) and reliable and clinically significant change (R+CSC) at each follow-up period, relative to pre-treatment levels (see Table S1). Reliable change cut-offs were calculated on the basis of the standard deviation of each measure at pre-assessment in the present sample and reliabilities (test-retest if available)

for each scale taken from the original scale development papers. Clinically significant change cut-offs were calculated using criterion c (participants are closer to comparison than clinical mean), using clinical group values taken from the pre-treatment assessment in the current sample and comparison group values taken from relevant scale development papers. Where comparison group values are not available, we use criterion a (falling more than two standard deviations away from clinical group mean in the direction of clinical improvement).

For the BADS (measuring behavioural activation), 70% of participants showed reliable improvement post-treatment (60% R+CSC), and these gains were largely sustained at each subsequent follow-up period. For the RAS-SF (measuring recovery orientation), 64% of participants showed reliable improvement post-treatment (36% R+CSC). There was a slight reduction at two-month follow-up, but these gains were then preserved at one-year follow-up (63% RC and R+CSC). For the ICECAP-A (measuring capability), 64% of participants showed reliable improvement post-treatment (55% R+CSC), with these gains being preserved at each follow-up period. A minority of participants showed reliable and clinically significant improvement on the RPA and ACS-P measures of appraisal style post-treatment, although at one-year follow-up there were promising signs of improvements in dampening appraisals (63% RC, 38% R+CSC).

Group level analysis

Complete data at all time points were available for ten out of eleven participants for the BADS, RPA and ACS-P and for nine out of eleven participants for the RAS-SF, and the ICECAP-A. As intended, there was a main effect of time for, BADS-total $F(3,27)=13.09$, $p<.001$, $\eta^2_p=.59$, RAS-SF total, $F(3,24)=17.46$, $p<.001$, $\eta^2_p=.69$, ICECAP-A tariff, $F(3,24)=5.16$, $p<.01$, $\eta^2_p=.39$, and RPA-dampening, $F(3,27)=11.43$, $p<.001$, $\eta^2_p=.56$. However, there was no significant effect of time for RPA emotion-focus, $F(3,27)=1.39$, $p=.27$, $\eta^2_p=.13$, RPA self-focus, $F(3,27)=2.73$, $p=.06$, $\eta^2_p=.23$, or ACS-P, $F<1$.

We resolved the main effect of time by running by pairwise comparisons (using least significant differences) for all measures where the main effect of time was significant. As intended, there was no significant difference between intake and pre-treatment for any measures, $Ps>.20$, with the exception that behavioural activation increased/decreased, $p=.048$. Also as expected, there was a significant improvement from pre to post treatment for all measures, $ps<.02$. There was significant improvement from pre to two-

month follow-up for a majority of measures, ps<.05, with the exception that the BADS change was only trend significant, p=.06. There were no significant differences between the two-month and one-year follow-up for any measures, ps>0.20.

We then went on to examine if gains were preserved at one-year follow-up, running paired sample t-tests between pre-treatment and one-year follow-up scores on each measure.

Gains, relative to pre-treatment, were maintained for RAS-SF, t=4.05, p<.01, ICECAP-A tariff, t=4.27, p<.01, and RPA dampening, t=4.64, p<.01. The BADS comparison showed a non-significant trend for improvements at one-year follow-up, t=2.03, p=.08. There continued to be no significant difference to pre-treatment levels on RPA-EF, t<1, RPA-SF, t=1.08, p=.32, or ACS-P, t<1.

Discussion

The behavioural activation, recovery and capability analyses mirror findings from the main manuscript, with robust and sustained improvement for a majority of participants. The ICECAP-A results, as far as we are aware, are the first indication that this scale is sensitive to change in depression. The results for the various appraisal measures were less clear cut. There were reductions in levels of dampening (particularly at one-year follow-up), but no changes in amplifying appraisals or fear of positive emotions. It is unclear if this reflects that ADepT does not target these constructs adequately or whether the measures are not sufficiently sensitive to change (at least when using the small sample size here that was only really powered to detect large effect sizes). This provides preliminary, partial support for the logic model of change underpinning ADepT. Further work is needed to test this logic model more fully.

S4.1: Topic Guide for Qualitative Interviews

1. Researcher to introduce themselves and explain the purpose of the interview.
2. Ask for consent to record- explain that this is so interviewer can give full attention without having to try and write it all down, and that recording will be transcribed and then analysed. "*We may quote your exact words when writing up the research but nobody will be able to identify you from them*…"
3. Participant to read and sign consent form

Opening question (an ice-breaker):

How did you find out about the research project?

Topic 1: Experience of Treatment

You have just received ADepT therapy. Could you tell me in your own words how you found this?

Probes (if required):

- Was it what you were expecting?
- Have you had therapy for depression before and, if so, how did it compare?
- Do you think this therapy focused more on reducing depression or building wellbeing?
- Did the therapy surprise you in any way?
- What was the best thing about it?
- What was the worst thing about it?
- Do you think there are any changes we could make that would improve your experience of the therapy?
- What did you think of the handouts and home exercises?
- Has therapy changed the way you think about your depression?

Topic 2: Therapeutic alliance

How did you get on with your therapist during ADepT therapy?

Probes (if required):

- How well do you feel your therapist related to you?
- Did you feel s/he understood you and wanted to help you?
- Did you feel that your therapist was actively involved in your treatment?
- Did you trust in your therapist?

Topic 3: Barriers to treatment

We are interested in things that help people fully engage with ADepT or prevent them from doing so. Please tell me what sort of things influenced how many sessions you attended and the amount of homework you completed during ADepT?

Probes (if required):

- If you dropped out of therapy before completing treatment, why was this?
- If you missed a session and then came back into treatment, what helped you to do this?
- Was the timing of sessions an issue for you at all?
- Was travel an issue for you at all?
- Were the views of friends and family an issue for you at all?
- Were other demands on your time an issue for you at all?
- Was your relationship with your therapist an issue for you at all?

Topic 4: Life after therapy

What are your views about likely impact of ADepT in the longer term?

Probes (if required)

- Preventing relapse/mood dropping
- Continuing to build wellbeing?
- What do you understand by the term wellbeing?
    - Job
    - Leisure
    - Self-Care
    - Relationships

Topic 5: Experience of research

How did you find taking part in the research study?

Probes (if required):

- What was your motivation to take part in a research study?
- How did you find completing weekly questionnaire measures?
- How did you find completing the questionnaire pack before and after treatment?
- How did you find completing the interviews before and after treatment?
- How did you find the volume of measures we asked you to complete?
- How did you find your contact with the research team (other than your therapist)?

Closing remarks:

Anything else you would like to say about ADepT/CBT or taking part in the research?

Debrief

1. Thank client for taking part.
2. Cover what will happen next.

3. Let participant know they can e-mail you with follow-up comments if they have any further thoughts, giving them a contact e-mail address.

S4.2: Qualitative Interview Thematic Analysis

We present the results broken down by higher level theme and then specific sub-theme.

Perceived benefits of treatment

Clients generally gave feedback on the treatment and felt that it had benefited them. For example:

- 'So it's given it a whole new angle for me for looking at wellbeing in depression. So for me it was very good' (AD03).
- 'I found it very helpful. I feel that there's improvement since before doing and after finishing. I'm just hoping that I can take it on with me now and try and keep going….' (AD05)
- 'It was just a really positive experience. I'd certainly recommend it to anybody that I knew was struggling with depression.' (AD08)
- 'So it was really well thought out, I think, and very well presented, and I have absolutely nothing to say in a negative way about it at all. …… I pray that it will become available to a wider source of people.' (AD11)

One client saw her depression as secondary to fibromyalgia. As the fibromyalgia was still present, her depression had not improved:

- 'It was all right, but because I've got fibro [myalgia] and my depression is from the pain, …I don't actually feel that I'm any less depressed from when I started …'[AD07]

One client had come in wanting concrete advice on a specific life issue which they felt the therapy did not address:

- 'I expected after therapy my biggest problem was gone, and I'm afraid after therapy my biggest problem is still here. I was thinking, right, when I get to the therapy they will tell me you go that direction [with regards to issue]. And they didn't tell me and I'm still thinking about it. Maybe I'm closer to go one direction, but I'm still not making that choice.' [AD10].

Views on wellbeing focus in ADepT

Focusing on wellbeing was not what clients had expected but was often experienced as helpful. For example:

- ‘I think I was dubious, not sure what it was all about. But …looking at all the different areas in your life that could be improved, where wellbeing could be improved, was definitely really, really good and perhaps something I wasn't expecting to look at’ (AD03)
- ‘We’ve moved from just one prong of the approach of dealing with the negatives, to dealing with the positives as well, and making them work for me. I think that is what’s been surprisingly good’ (AD06)
- ‘I think the idea of helping you to rediscover pleasure and enjoyment and confidence in daily life, helping you reconnect and feel less alienated. And that in itself reduced the depression massively’ (AD08)

Clients commented on how the positive future focus of ADepT was useful, although there were mixed views about whether there was also a need to focus on the past. For example:

- ‘It was a lot, lot better than counselling because of the way that it focuses on looking forwards and where you want to get to and how to get you there. I’ve found that much more helpful than other talking therapies and counselling, where you look backwards at things that have happened in your life and try to understand why you’re having bad reactions….that made we dwell on bad things rather than moving forward with them’ (AD12)

Improvements in wellbeing were seen as having a knock on effect of depression. For example:

- ‘I think it focused more on wellbeing and I think reducing depression as a result of focusing on wellbeing’ (AD08)
- ‘I think it was more on building my wellbeing, and the knock-on effect was that the symptoms of depression lessened’ (AD12)

<u>Changing relationship to depression during ADepT</u>

Clients reported no longer seeing depression as ‘the whole of them’ and instead seeing it as something that came over them. For example,:

- ‘I think I assumed depression was the whole of me, whereas this therapy has taught me it’s just part of me’ (AD03)

- ‘Now I consider myself as a person who has depression and anxiety, rather than a depressed and anxious person.’ (AD12)

They also described accepting their depression more: this is something that may recur and that they did not to beat themselves up about this. For example:

- I'm, unfortunately, confident that I'm going to have these problems pretty much for the rest of my life in one form or another, but it comes back to dealing with the problems and not letting them get too bad …. having that written document saying when I start noticing this with my mood, I will do this, that, whatever it is that is needed, is something that I think will be useful.’(AD08)
- ‘I learned ... there is nothing wrong with me having bad feelings or bad thinking about something. Other people have it as well... its not only me… its normal’ (AD10)

They reported feeling more resilient to challenges and setbacks in life, in part by recognising low mood will pass and noticing the overall positive direction of travel rather than catastrophising about the setback. For example:

- ‘My depression is still here, I am not coping with some things as well as I would like, I’m finding the setbacks in life knock we down harder than they should do, but the difference between six months ago is that I’m also bouncing back quicker as well’ (AD06)
- ‘and if I had a difficult week, then I flagged it up as being a difficult week rather than the end of the world…it helped me to see that my mood was going up and down but that the projection was generally going up’ (AD08)
- ‘It’s different because before I’d just dwell on the …. one tiny little thing that went wrong that I thought was a big problem’ (AD09)

Clients reported feeling more in control of their depression, spotting drops in mood early and having new tools to manage it. For example:

- ‘… that it doesn’t have to control me, which it has done for the last however many years… before it was really overwhelming, whereas now I know that it hasn’t got to overwhelm me if I take the right steps’ (AD03)

- ‘It’s helping me to look at it as something to be handled, rather than something I’m not completely out of control with’ (AD05)
- ‘I feel better, like I can catch it more, recognise and catch it before it starts to have a negative impact…’ (AD08).
- ‘It was made clear as well that once you've had periods of depression in the past it's likely that you are going to begin to have periods of depression again, the black dog. But I believe now that I've got the tools that I don't have to fall right down that hole, I can actually look at ways to stop that happening at a much earlier stage.’ (AD11)
- ‘ If I have worries or episodes of low mood, then I can think about them more rationally and they don’t take over my mind’ (AD12)

Importance of therapeutic relationship in ADepT

Clients valued a non-judgemental, caring, sensitive listener they connected with, was trustworthy, was able to ‘contain’ them, that put them at their ease, and was at the same time professional. For example:

- ‘It was good having someone to finally listen. Feeling in a relaxed way you were able to open up and get some of these things off your chest …. I felt comfortable and in a position where I was able to speak freely without being judged….. Couldn’t have been nicer to me really’ (AD03)
- ‘ I got on well with him [therapist]. He’s very helpful and really patient…..He’s a really nice guy….He’s still very down with the kids when you talk to him….it just felt like a person-to-person conversation. There wasn’t a barrier of sorts, given our differences in experience.’ (AD05)
- ‘[Therapist name] was a lovely person. She managed to blend being friendly and approachable with her professionalism very well. We always managed to be friendly, despite the fact that we’re discussing quite unpleasant topics at times. ….. it’s a very genuine interest she was taking in my case …..’ (AD06)
- ‘I was really impressed with the professionalism and care I received, and the patience… It felt very human ……The counsellor was just a really good listener. He would prompt and probe issues that reveal sensitivity. [He was] non-judgemental and helped you to see the answers to certain things by

making you think you'd arrived therefore yourself really, even though he'd probably steered you there all the way along. ……having somebody who I felt really cared but wasn't involved in my own personal issues was really helpful… to have a dispassionate but compassionate person to talk to …' (AD08).

- 'I think I felt comfortable with the person and I felt like she [the therapist] could understand me…… I think when you get here [to therapy] and you feel that connection with the therapist…someone to talk to… it gives you strength' (AD10)

Clients appreciated therapists bringing (an appropriate amount) about themselves into the work and being willing to engage in broader dialogue. For example:

- '…… He certainly didn't share any personal stuff about himself, but he brought in enough about him as a person or his life to allow me to feel that there was empathy there.'(AD11)
- 'Obviously therapists are professionals who are there to help you. But especially at the beginning I think I just found it more relaxing to feel they were slightly less of a therapist and sort of had a fertile sessions talking about random stuff to get to know them a little bit before you start opening your heart to them' (AD12)

At the same time, some clients acknowledged that it was hard initially to trust the therapist, they wondered if the relationship was 'genuine', and worried about being judged. This sense of trust grew over time. For example:

- 'Probably lots of people with my kind of problems have issues trusting people. So it's a difficult thing to be asked to trust this person, even when you know they are a professional… Sometimes I felt that I was irritating/boring my therapist. But I think that was probably my reading of the situation, rather than how it actually was ….he [the therapist] never said anything that implied I wasn't working hard enough or achieving the right things, which is why I think it was my own perception of it. …we've developed a really good relationship, but it took me a bit of time to get used to talking to him and trusting him [the therapist]' (AD12)

<u>Helpful aspects of ADepT</u>

Identifying what was important to them and living life in accordance with these values was found to be particularly useful. For example:

- 'It's the first time I have really started to look at what I would like for life. So hopefully I'm more positively selfish from now on. There's a difference between being actually selfish …. It's so easy to get trapped in putting everybody else first and then just completely forgetting about you' (AD05)
- 'The focusing on values I found very, very helpful… a lot of what I was doing, I think, was not that in line with my values, and so to reorientate my life and work practice back in that way was really helpful' (AD08)
- 'And it was great for me to be able to relook at the values I thought I held and actually to be able to change them to the person I am now, and to be able to get rid of some old habits … I was surprised, the values in all the different quadrants of my life as we were working just changed. What is really important to me isn't what perhaps I thought it was that's been inculcated by work ethic. I'm a striver, I'm an achiever, and I don't need to be that anymore' (AD11)

Clients commented that engaging with everyday wellbeing activities was useful but that enjoying the moment more was still a work in progress. For example:

- 'Just stopping to take 5 minutes and say I am actually enjoying sitting here, having a nice cup of coffee, with my wife, for example. It's not that I didn't appreciate moment moments before: I did, but they got lost in the depression. Bringing that a bit more into the foreground has been very, very useful' (AD06)

Clients reported that an emphasis on self-care and self-compassion had been beneficial, although at times they still struggled with a sense they did not deserve this. For example:

- 'I'd got out of the habit of looking after myself very well. So I particularly wasn't eating very well and I was smoking too much … So I think what it's done is it's made me value myself more and realise that my health has a direct impact on my mental state…' (AD08).

Clients reported increasingly recognising the impact of depression on their degree of social contact, the positive impact of social contact on their mood, and the value in seeking support from others. For example:

- ‘The illness has probably cost me a couple of friends. The flip side of that is I’m getting to the point where I’ve started making new social contacts as well. … its [ADepT] enabled me to get into a place where I’m mentally strong enough to start forming new relationships with people.’ (AD06)
- ‘Before I thought I was alone in depression, I didn’t see any way that anybody else could help me, I put up walls, I became very insular, didn’t want anyone near me, didn’t want to talk to people. Now I’ve realised that its possible to involve other people and allow them to help in a way that I can move forward when I am depressed.’ (AD11)
- ‘I've worked out with my therapist that if my mood starts to drop … I then isolate myself, because I didn't want to be a boring, miserable person in the group. And I've worked out that that wasn't helping my mood to improve. So if I feel my mood dropping then it's important to maintain social contact, because that's one of the best ways for me of bouncing the mood back up.’ (AD12).

Clients reported liking the regular structure and practical focus. For example:

- [I liked] how practical it was. I got tasks to complete every week. Whether that was something on paper or an active going out and doing something, there was always some concrete milestone to hit between sessions. And that made it easier to track progress, and you could feel like you were achieving something by doing it …in previous therapies I've never really felt that I was moving anywhere.’ (AD12)

<u>The challenge of engaging with ADepT</u>

There was a recognition that ADepT asked a lot from clients, in terms of completing homework and understanding sometimes complex material. While it was acknowledged that ADepT exercises and homework were valuable, it was also hard to do when depressed and failure to do so could trigger self-criticism or anxiety. For example:

- ‘[Handouts and homework were] really good, really useful. I think if I hadn't had those, I probably wouldn't have got anything like as much out of the sessions. ……because things are all a bit hectic in my life at the moment, because I'm involved in various projects which take up a lot of time….I was thinking sometimes, oh God, I've got to do this as well. But actually it was really valuable, because by doing it, it helped me to get everything else into perspective.’ (AD08)

- ‘I started off putting it off. But as we got a bit better in the thinking, they weren't so much a case of I've got to do that. It felt that it was actually starting to be helpful.’ (AD09)
- ‘I think probably in terms of volume, but also in terms of understanding the quick-fire way in which it's put. A lot was put into an hour. And I think you needed to be quite understanding cognitively to be able to take on board everything that happened in that hour…. I think if I'd been in a deep or deeper depression, I might have thought I can't do that and panicked.’ (AD11)

Views on ADepT handouts and homework exercises

Clients felt the handouts helped them consolidate their learning and remember more and would be a valuable resource in the future. For example:

- ‘So all [the] bits and pieces each week to read is just something else to add to like the book. And then if I'm feeling stressed out or something, I can like read it and just bring myself back. So that's really helpful..... it's hopefully just something that I can climb back on the horse with if I start to slip.’ (AD05)

Clients reported finding many of the exercises were useful. For example:

- ‘The task where I looked at the bullseye, finding the sweet spot between vocation and leisure and relationships and self-care. That was really helpful to make me see how out of kilter I was…’(AD08).
- ‘The last couple of sessions that we did with [therapist] were focused on relapse prevention. There was a document that we put together, a relapse prevention sheet, of making sure that you notice the signs and take them seriously so that as soon as your mood starts dipping you put in place the actions that are needed to prevent the slide from getting any worse. That's something that I'll be keeping to hand.’ (AD08)
- ‘[the mapping tool] taught me to see the negative thoughts and to try and turn them around before they become a problem. In some ways it's like trying to do the opposite of what the instinct or the first thought is. It's, no, I'm not going to enjoy that so I won't do it, instead of trying to think, well, it probably won't be that bad, so let's do it. It probably helped me realise that I wasn't doing quite as

badly as I thought. So, helped me to see the positives, rather than just focus on the negatives or the downsides.' [AD09].

- '[views on mapping tool]…. the way we would look at a certain situation and the way the old me would approach it, and the way that perhaps it would be better for me to approach it now. That was very useful.' (AD11)

Not all clients experienced all exercises as helpful and some clients reported they did not fully engage with all exercises. For example:

- '[in relation to mood diary] There was a diary you had to fill in about what you were doing each day hour by hour. That's the only thing that I found sort of a bit tedious.' (AD03)
- 'I've tried to do everything I've been told, but he said about writing something about something wonderful, or what gives you pleasure. Well, I never actually picked that bit of paper up again.' (AD07)

Views on booster sessions

Clients generally were positive about being offered booster sessions (even if they chose not to use all five). For example:

- 'I had a monthly follow-up with (counsellor) last week, which, again, was really positive.' (AD08)
- 'I'm very pleased that there's a few opportunities to follow up, I think that's important for me. … I'm aware that there's somebody there … if I feel that I'm not able to use the tools that I've been given.' (AD11)

Areas to improve in ADepT

Clients accurately identified that some of the materials could be more refined and that the therapists were learning a new protocol so were not fully fluent in early sessions. This meant that not all materials introduced were returned to in subsequent sessions:

- As I understand it, this is quite early in the study. Some things will become a little bit more polished, a little bit more practised, if you will.' (AD06)
- '….[some materials] I couldn't get my head around …. They just didn't seem to be mentioned again' (AD07)

Reflecting the importance of the social domain to clients, it was recommended there should be an even greater focus on social activation:

- 'My own experience with depression …. it's really isolated me. Maybe a little bit more focus on reintegrating yourself with your social circle and friends and family might be helpful.' (AD06)

One client reflected that some of the ground covered was not new to them, but recognised it was important to include as not all clients will have had prior treatment:

- 'Some of the coping techniques and stuff have been unnecessary, because I already know them, I'm very familiar with them. But somebody coming to this service new, you know, their first proper treatment for depression, they may well find that useful.' (AD06)

It was felt that more attention could have been paid to how life stage is impacting on depression, in one example around the onset of the menopause:

- 'I think a couple of times that there was very little attention paid to outside factors. for instance, I'm of an age where I'm just starting to enter into the menopause and my hormones are all over the shop, and I think that's having a bit of an impact on my mood…..I just wonder whether there would be a little bit of scope to look at how people's life stage might be affecting their depression as well.' (AD08)

<u>Barriers to engaging with treatment</u>

Some clients initially reported being sceptical that therapy could help, given previous unsuccessful experiences with other treatments, but that they were glad they persisted. For example:

- 'I found it really helpful. At first I was a little bit resistant to the sort of tick box type style of it. But actually the more I met with [therapist] the more I realised it was a really helpful way of creating a framework and a set of tools that you could keep referring back to, to help you to help yourself. So I did find it really helpful ……' (AD08)
- 'I'm very pleased with the way it worked out. Initially I was quite sceptical, because talking therapies in the past just haven't worked, they haven't solved the problems that I've had, haven't even made any real progress. I agreed to be referred here because I was getting to the point where I was

ill enough to try anything. ….So, I was like I say, prepared to give anything a go, but I wasn't holding much hope for it.' (AD08)

A number of clients reported that symptoms of depression, anxiety and pain also made it difficult to attend sessions. For example:

- 'There were a couple of times where I felt early on in the therapy that I wasn't sure that I was strong enough to come in and do it, because anxiety has got me shunning social contact a lot in the past …. there have been times where I've thought I know it will be good for me but I'm not sure I can face going to it.' (AD06)
- 'A lot of weeks I used to think, I don't know if I can be bothered to go. But that's what the fibro [myalgia] makes me feel like, I can't be bothered to get myself out of the house. But I used to come up here and feel a lot better.' (AD07)
- 'There was one week where I really didn't feel like coming at all, depression-wise my levels of energy were right down. But I did come and I felt so much better.' (AD08)

What helped clients to attend at these times despite low mood was a sense sessions were being beneficial. For example:

- 'Because this is so helpful, I was determined to come in every week. But given my condition [depression], you wake up in the morning and you just don't want to do anything But I would still push through that and come in.' (AD05)
- 'Knowing that every time I've come before, I've come away feeling a whole lot better [helped me keep coming]' (AD08).

Practical life issues also got in the way of attendance. For example:

- 'There was one time where I missed it accidentally because my alarm didn't go off…..I was beating myself up over that next week, but I still got up and came in, 'cos I didn't want to quit.' (AD05)
- '…just life events really. There were a couple of weeks where I had like car problems, there was one week where I had a commitment to a friend to help out with something.' (AD08)

Text message reminders and therapists supporting them around attendance was helpful to keep them engaged in treatment. For example:

- ‘Initially I said because I wasn't very well I wasn't sure if I wanted to continue with treatment. And I had an email from the therapist to say that he understood it was a difficult time, but it's often at these times that therapy's beneficial. So that really sort of helped me to come back on track.’ (AD03)
- ‘The text reminders have been really helpful …’ (AD08)

References

Al-Janabi, H., Flynn, T. N., & Coast, J. (2012). Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. *Quality of Life Research, 21*, 167-176.

Al-Janabi, H., Flynn, T. N., Peters, T. J., Bryan, S., & Coast, J. (2015). Test–retest reliability of capability measurement in the UK general population. *Health Economics, 24*, 625-630.

Cook, J. A., Copeland, M. E., Floyd, C. B., Jonikas, J. A., Hamilton, M. M., Razzano, L., ... & Boyd, S. (2012). A randomized controlled trial of effects of Wellness Recovery Action Planning on depression, anxiety, and recovery. *Psychiatric Services, 63*, 541-547.

Corrigan, P. W., Salzer, M., Ralph, R. O., Sangster, Y., & Keck, L. (2004). Examining the factor structure of the recovery assessment scale. *Schizophrenia Bulletin, 30*, 1035-1041.

Feldman, G. C., Joormann, J., & Johnson, S. L. (2008). Responses to positive affect: A self-report measure of rumination and dampening. *Cognitive Therapy and Research, 32*, 507-525

Flynn, T. N., Huynh, E., Peters, T. J., Al-Janabi, H., Clemens, S., Moody, A., & Coast, J. (2015). Scoring the ICECAP-A capability instrument. Estimation of a UK general population tariff. *Health Economics, 24*, 258-269.

Hayes, A. (2015). Facilitating emotional processing in depression: The application of exposure principles. *Current Opinion in Psychology, 1*, 61-66.

Kanter, J. W., Mulick, P. S., Busch, A. M., Berlin, K. S., & Martell, C. R. (2007). The Behavioral Activation for Depression Scale (BADS): psychometric properties and factor structure. *Journal of Psychopathology and Behavioral Assessment, 29*, 191-202.

Keeley, T., Coast, J., Nicholls, E., Foster, N. E., Jowett, S. & Al-Janabi, H. (2016). An analysis of complementarity of ICECAP-A and EQ-5D-3L in an adult population of patients with knee pain. *Health and Quality of Life Outcomes, 14*, 36.

Kocalevent, R., Hinz, A., & Brahler, E. (2013). Standardization of the depression screener Patient Health Questionnaire (PHQ-9) in the general population. *General Hospital Psychiatry, 35*, 551-555.

Lowe, B., Decker, O., Muller, S., Brahler, E., Schelberg, D., Herzog, W., & Herzberg, P. Y. (2008). Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the General Population. *Medical Care, 46*, 266-274.

Melka, S. E., Lancaster, S. L., Bryant, A. R., Rodriguez, B. F., & Weston, R. (2011). An exploratory and confirmatory factor analysis of the Affective Control Scale in an undergraduate sample. *Journal of Psychopathology and Behavioral Assessment, 33,* 501-513.

Mitchell, P. M., Al-Janabi, H., Byford, S., Kuyken, W., Richardson, J., Iezzi, A., & Coast, J. (2017). Assessing the validity of the ICECAP-A capability measure for adults with depression. *BMC Psychiatry, 17*: 46.

Plummer, E., Manea, L., Trepel, D., & McMillan, D. (2016). Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic meta-analysis. *General Hospital Psychiatry, 39,* 24-31.

Raes, F., Daems, K., Feldman, G. C., Johnson, S. L., & Van Gucht, D. (2009). A psychometric evaluation of the Dutch version of the responses to positive affect questionnaire. *Psychologica Belgica, 49,* 1-18.

Williams, K. E., Chambless, D. L., & Ahrens, A.H. (1997). Are emotions frightening? An extension of the fear of fear concept. *Behaviour Research and Therapy, 35,* 239-248.

Table S1

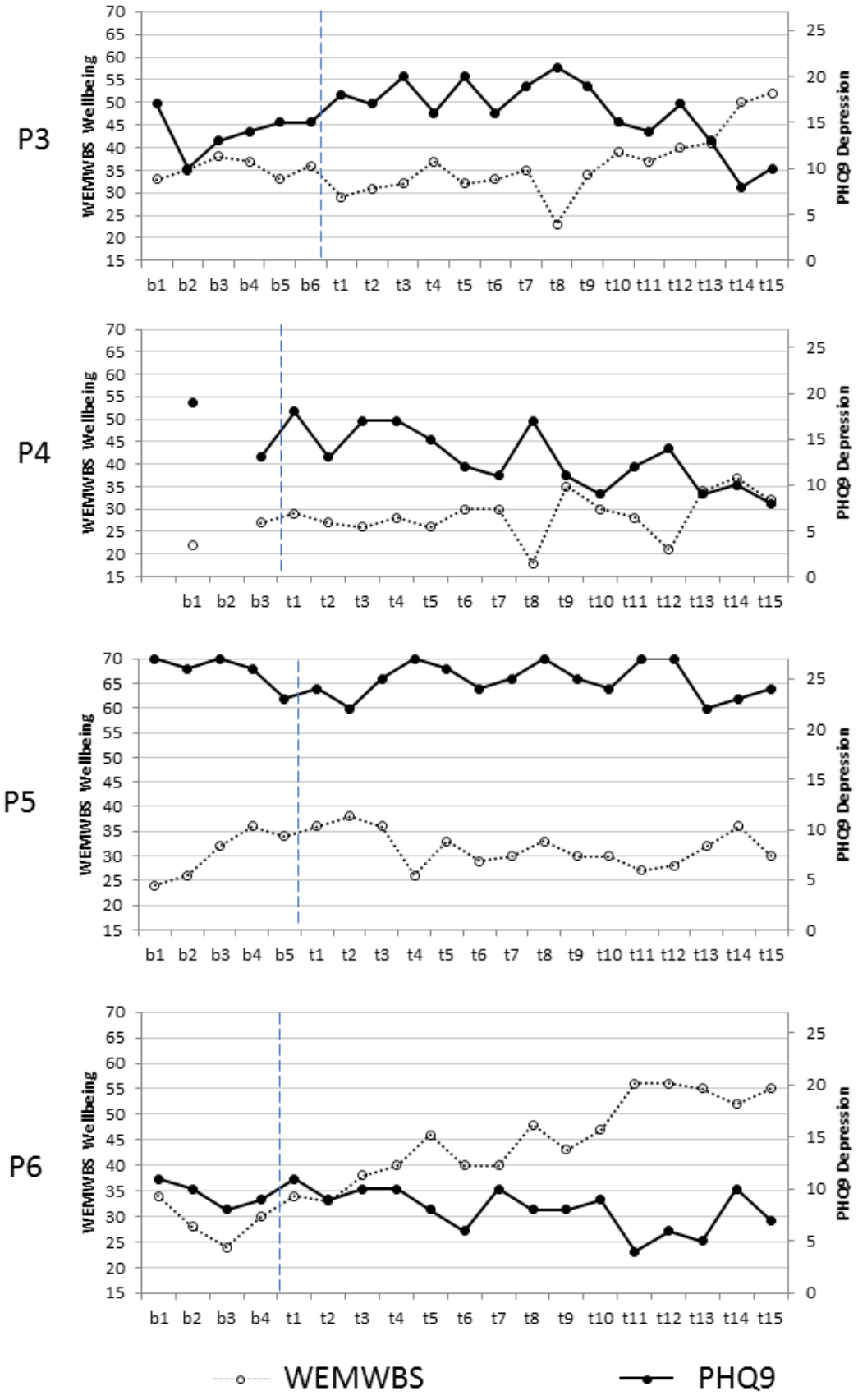
Anxiety, functioning, behavioural activation, recovery and appraisal style, and capability outcomes at each phase (and proportion showing reliable and clinically significant change)

	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	Mean (SD)	RC (CSC)
BADS total													
Intake	62	65	64	70	76	65	73	80	108	93	77	76.80 (14.26)	-
Pre	69	50	110	77	86	63	75	83	115	98	89	83.20 (19.55)	-
Post	90*	-	92	92	107**	118**	121**	91	161**	132**	116**	112.00 (22.81)	70% (60%)
2mFU	73	-	63+	130**	85	106**	99*	99	161**	139**	113**	106.78 (30.24)	60% (50%)
1yrFU	99*	-	72+	114**	99	143**	-	87	150**	131**	-	111.88 (27.60)	63% (50%)
RAS-SF total													
Intake	57	51	59	67	56	56	63	45	69	67	58	58.82 (7.24)	-
Pre	57	46	63	60	59	57	72	63	57	64	60	59.75 (6.38)	-
Post	73*	-	63	77**	69*	78**	77	64	84**	81**	71*	73.66 (6.83)	64% (36%)
2mFU	59	-	-	85**	67	79**	76	62	84**	85**	60	72.98 (10.94)	40% (40%)
1yrFU	82**	-	58	84**	57	85**	-	56+	95**	89**	-	75.83 (16.03)	63% (63%)
ICECAP-A													
Intake	.53	.37	.57	.54	.59	.61	.39	.25	-	.61	.51	0.50 (0.12)	-
Pre	.44	.47	.53	.44	.48	.53	.51	.15	.34	.69	.51	0.46 (0.13)	-
Post	.56	-	.36+	.76**	.77**	.87**	.76**	.32*	.97**	.92**	.51	0.68 (0.22)	64% (55%)
2mFU	.34	-	.47	.69**	.71**	.89**	.51	.44*	.98**	.83**	.58	0.64 (0.21)	55% (45%)
1yrFU	.80**	-	.64	.73**	.56	.89**	-	.36*	.89**	.76	-	0.70 (0.18)	63% (50%)
RPA-D													
Intake	23	21	29	18	26	22	28	24	27	18	21	23.36 (3.80)	-
Pre	26	26	32	17	26	16	25	18	22	20	19	22.45 (4.95)	-
Post	22	-	32	16	17**	11	18**	19	18	15	12**	18.00 (5.89)	27% (27%)
2mFU	24	-	30	8**	18**	16	21	18	18	12**	16	18.10 (6.09)	27% (27%)
1yrFU	19*	-	24*	17	13**	13	-	13	11**	12**	-	15.25 (4.43)	63% (38%)

RPA-EF													
Intake	9	5	10	7	8	7	9	7	17	15	9	9.36 (3.59)	-
Pre	8	7	11	6	12	6	13	10	6	18	8	9.45 (3.75)	-
Post	13	-	16	9	7	10	16	10	18**	14	7	12.00 (3.94)	9% (9%)
2mFU	15**	-	7	8	13	10	14	7	18**	16	5	11.30 (4.47)	18% (18%)
1yrFU	16**	-	6	9	7	10	-	10	16**	14	-	11.00 (3.89)	25% (25%)
RPA-SF													
Intake	5	4	6	4	7	4	7	4	14	7	7	6.27 (2.90)	-
Pre	6	5	8	4	7	5	8	6	8	9	6	6.55 (1.57)	-
Post	11**	-	6	5	7	9**	10	6	11**	9	8	8.20 (2.15)	27% (27%)
2mFU	10**	-	4+	7*	10**	6	9	5	11**	10	7	7.90 (2.42)	36% (27%)
1yrFU	9**	-	4+	7*	5	9**	-	7	12**	9	7	7.67 (2.40)	50% (38%)
ACS-P													
Intake	28	22	38	17	23	33	36	24	39	41	32	30.27 (7.98)	-
Pre	20	25	43	8	26	31	32	22	44	40	34	29.50 (10.87)	-
Post	30	-	49	15	20	27	34	20	37	34	32	29.80 (9.89)	0% (0%)
2mFU	25	-	53	19+	11**	37	31	18	37	35	33	29.90 (12.08)	9% (9%)
1yrFU	30	-	46	37+	12**	30	-	24	20**	29*	-	28.50 (10.34)	38% (25%)

Supplementary Figure 1a

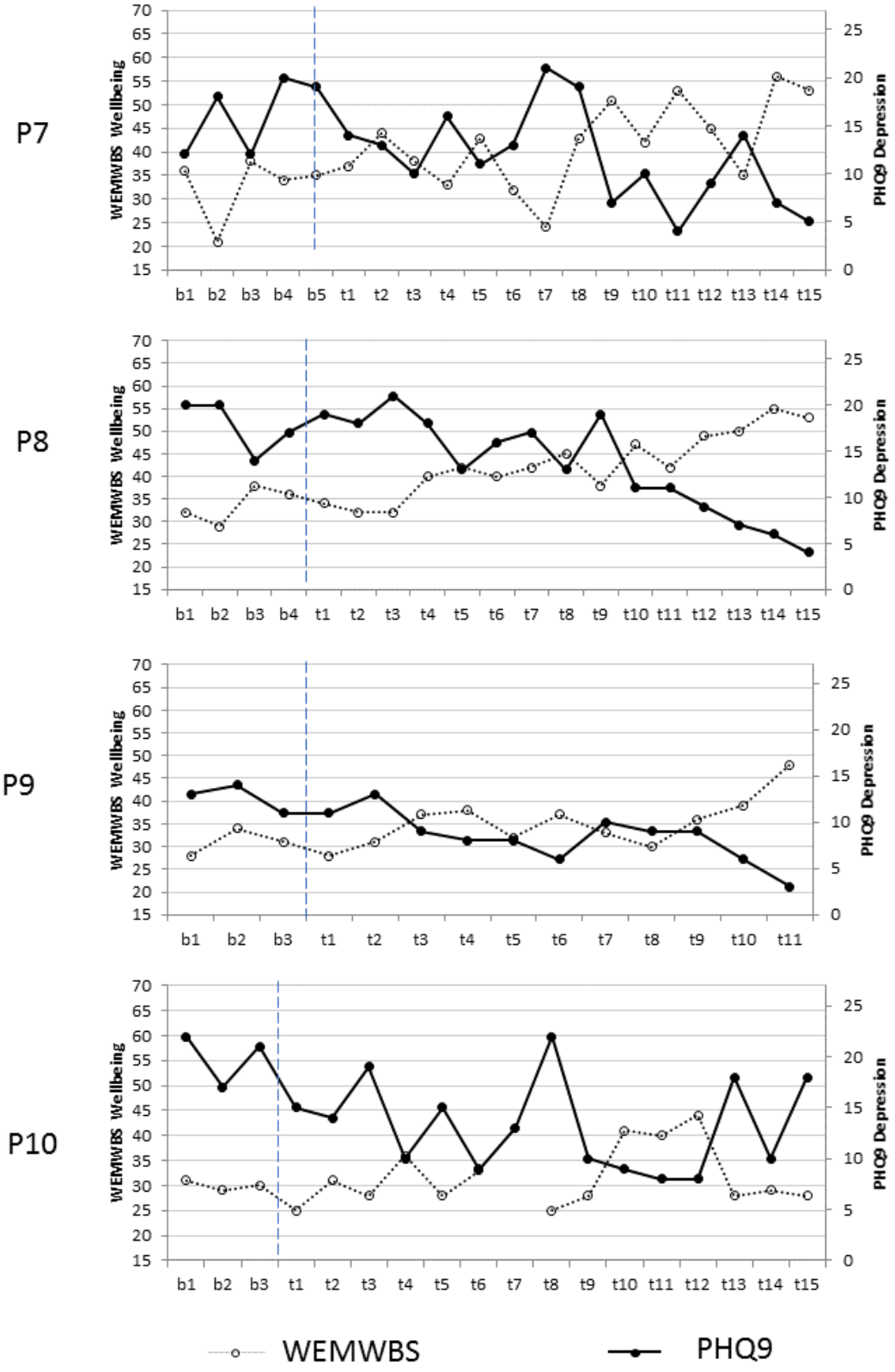
Depression and wellbeing weekly levels for each participant during baseline and treatment phase for participants 3-6



Note- b = baseline sessions, t = treatment sessions, vertical dotted line indicates start of acute treatment

Supplementary Figure 1b

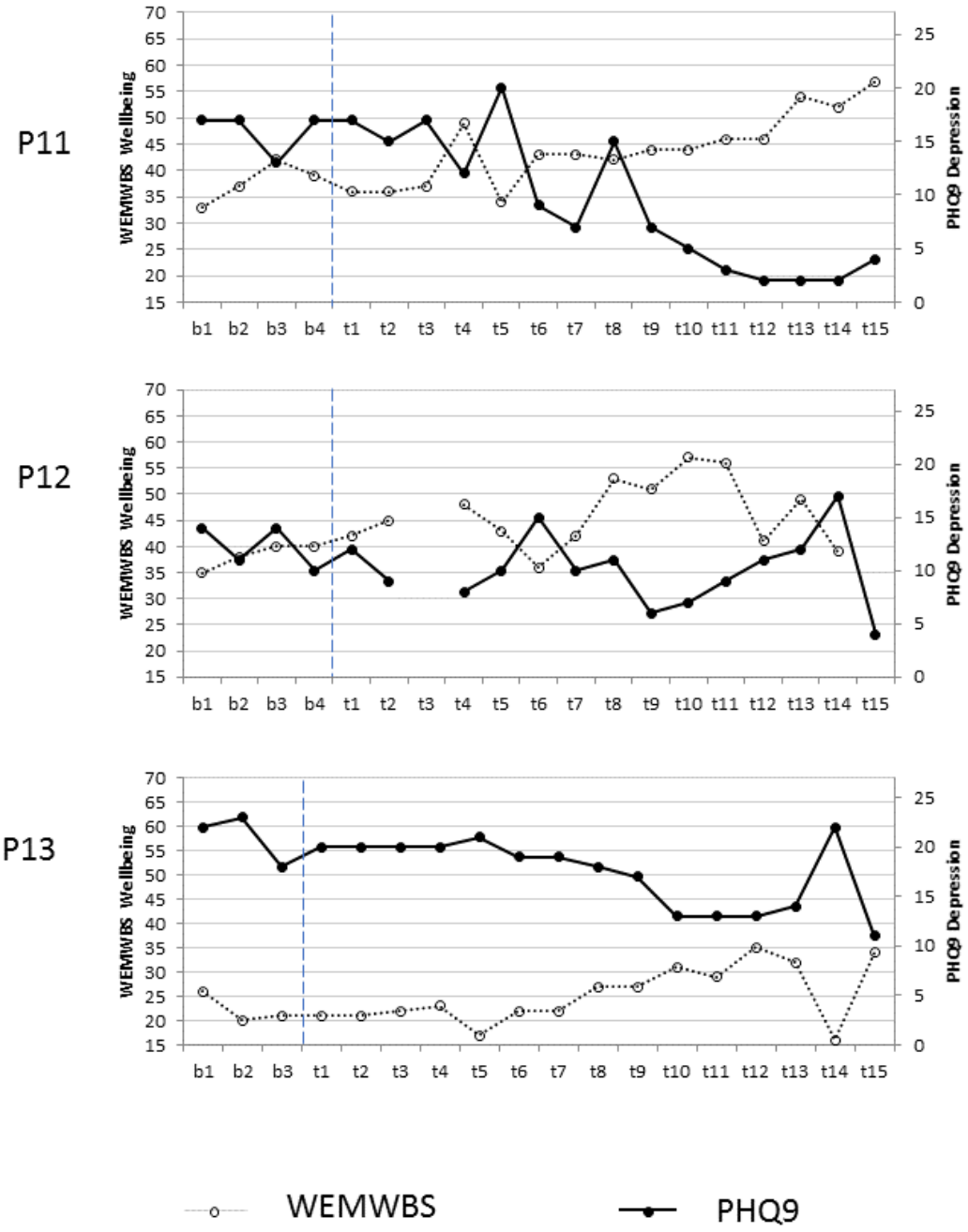
Depression and wellbeing weekly levels for each participant during baseline and treatment phase for participants 7-10



Note- b = baseline sessions, t = treatment sessions, vertical dotted line indicates start of acute treatment

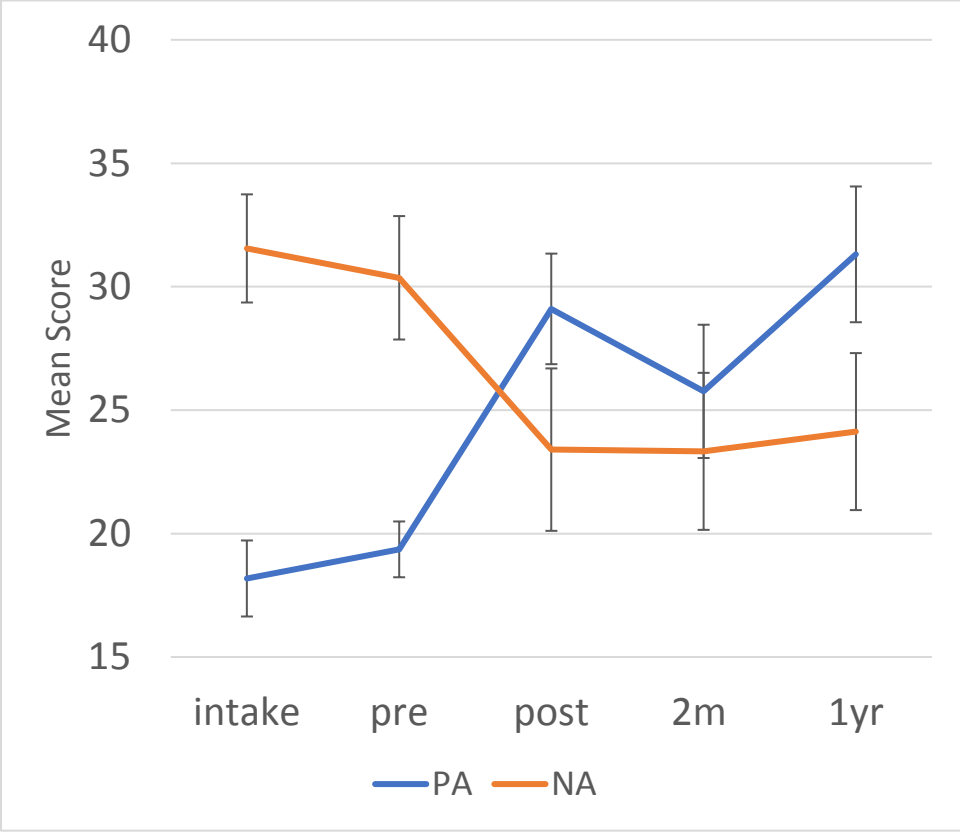
Supplementary Figure 1c

Depression and wellbeing weekly levels for each participant during baseline and treatment phase for participants 11-13



Note- b = baseline sessions, t = treatment sessions, vertical dotted line indicates start of acute treatment

Supplementary Figure 2
Mean score for PANAS positive and negative affect at each assessment point



Note – data are mean (standard deviation) values.