



# Cognitive Therapy or Behavioral Activation for Major Depressive Disorder in Dutch Mental Health Care: Pilot Effectiveness and Process Trial

Felix Bolinski<sup>1</sup> · Gert-Jan Hendriks<sup>2,3,4</sup> · Seline Bardoel<sup>5</sup> · Steven D. Hollon<sup>6</sup> · Christopher Martell<sup>7</sup> · Marcus J. H. Huibers<sup>1,8</sup>

Published online: 28 September 2018

© The Author(s) 2018

## Abstract

Few studies have investigated the association between psychotherapeutic approaches and their processes in routine psychological practice. We compared cognitive therapy (CT) and behavioral activation (BA) on their effectiveness and change processes. Forty-three depressed patients participated in our trial. Scores on depression and specific and nonspecific factors were collected at seven time points and analyzed using RM-ANOVA and multiple linear regressions. No differences in depression reduction emerged between conditions. Most processes changed during therapy. Only measures of negative cognitions and behavioral activation interacted with treatment condition. Change on the processes did not predict symptom alleviation. Similarly, reductions in depression were not followed by change on any of the process measures. Both psychotherapeutic approaches led to significant and comparable symptom reduction. There was no clear evidence of differential change with respect to purported underlying mechanisms. The results are discussed in the context of therapy and research.

**Keywords** Depression · Psychotherapy · Cognitive therapy · Behavioral activation · Change mechanisms

The high prevalence of depression and the burden it places on the individual and society has long been recognized (Cuijpers et al. 2007; Fried and Nesse 2014; Greenberg, Fournier, Sisitsky, Pike, and Kessler 2015; Kessler et al. 2005; Mathers, Fat, and Boerma 2008; Mehta, Mittal, and Swami 2014; Trivedi 2004). Next to antidepressant medication (ADM), several psychotherapeutic approaches to the treatment of major depression exist, among which cognitive therapy (CT; Beck 1964) is one of the most extensively researched (Cuijpers et al., 2013b; Hofmann, Asnaani, Vonk, Sawyer, and Fang 2012). Evidence has accumulated showing that CT can be as

---

✉ Felix Bolinski  
f.bolinski@vu.nl

effective as ADM (DeRubeis, Gelfand, Tang, and Simons 1999). DeRubeis et al. (2005) conducted one of the first large trials in which they compared ADM, CT, and a pill placebo for patients with moderate to severe depression. Their overall results did not show significant differences between the active conditions at the end of treatment. However, both outperformed the pill placebo group. A meta-analysis by Cuijpers et al. (2013a) points in a similar direction. The authors found no significant differences between CT during the acute phase and continued ADM. This held for both short- and long-term effects. CT is now recommended as a first line choice in the treatment guidelines for depression in the Netherlands (Spijker et al. 2013).

The study reported herein is a pragmatic pilot trial that has been conducted as part of a series of Dutch trials in outpatient settings. Despite the original CT manual (Beck, Rush, Shaw, and Emery 1979) recommending two weekly sessions in the beginning of therapy, most therapists in Europe generally see their patients only once per week (Bruijniks et al. 2015). Therefore, the first aim of this trial was to test the implementation of a regimen of two psychotherapy sessions per week. Secondly, we wanted to inform a large ongoing study described in Bruijniks et al. (2015) on the effects of session frequency and other mechanism of change in psychotherapy for depression. Thirdly, we intended to provide a benchmark for the effectiveness and mechanism trial by Lemmens et al. (2015), which compared one session per week of CT to interpersonal psychotherapy (IPT). In addition to session frequency, we investigated specific and nonspecific processes that are believed to underlie symptom change. To do this, we compared a group of patients receiving CT to a group receiving behavioral activation (BA), while also measuring factors presumed to underlie symptom change in depression at multiple time points, in an attempt to understand how these treatments work. BA was chosen as a comparator due to its ease of application and renewed interest in the approach (Richards et al. 2016). Earlier research has indicated that BA is as effective as CT (Barth et al. 2013; Cuijpers, Andersson, Donker, and van Straten 2011; Dimidjian et al. 2006; Jacobson et al. 1996).

With respect to the purported mechanisms of change, CT posits that resolution of negative cognitions and schemas leads to improvement (Beck et al. 1979; Butler and Beck 1995), an assumption that has received considerable validation (DeRubeis et al. 1990; Renner, Lobbestael, Peeters, Arntz, and Huibers 2012; Warmerdam, van Straten, Jongasma, Twisk, and Cuijpers 2010). However, it has been shown that such cognitions co-vary with the disorder (Beavers and Miller 2005; Teasdale 1983). Moreover, though CT places unique importance on negative thoughts, a reduction in such cognitions seems to accompany most interventions (Fava, Bless, Otto, Pava, and Rosenbaum 1994; Garratt, Ingram, Rand, and Sawalani 2007; Kovacs, Rush, Beck, and Hollon 1981), making it difficult to clearly identify them as the underlying change mechanism. An analysis of the temporal sequencing of the outcome (e.g., depression) and the potential mechanism or mediator (e.g., negative cognitions) of a specific treatment approach is necessary (Kraemer, Wilson, Fairburn, and Agras 2002).

In contrast to CT, the BA model proposed by Lewinsohn (1974) makes no assumption about clients' thinking patterns. Instead, it presumes a lack in positive reinforcement as the maintaining factor of depression, and emphasizes the importance of behavioral activation. Therefore, we assessed this latter factor as well.

Finally, the therapeutic relationship has long been hypothesized as a nonspecific driving force in therapy success (Horvath and Symonds 1991; Martin, Garske, and Davis 2000); we therefore also assessed participants' perception of the relationship with their therapist. To our knowledge, this is the first study that directly compared CT and BA in a Dutch healthcare setting. In this report, we provide the results of the pragmatic pilot trial.

## Method

### Design

Participants were recruited from a Dutch outpatient treatment center (Hendriks & Roosenboom, Arnhem; now Dr. Bosman, Arnhem). They provided informed consent before being included in the study. Inclusion criteria were a diagnosis of depression based on the results of the Structured Clinical Interview for DSM Disorders (SCID-I; First, Spitzer, Gibbon, and Williams 2002) and a score above 20 on the Beck Depression Inventory, Second Edition (BDI-II; Beck, Steer, Ball, and Ranieri 1996). Participants were excluded if they had severe comorbid conditions, such as substance dependence or schizophrenia. Moreover, participants were also excluded if they had received CT or BA within 12 months preceding the study, were taking antidepressant medication that could not be discontinued, or reported any level of suicidal ideation. The last criterion was applied due to the specific care that would be called for, in particular, the need for medication. Since the treatment center only provides specialist care for anxiety and mood disorders, incoming patients with (severe) personality or bipolar disorders are always referred elsewhere. Eligible patients were allocated to therapists who were in turn randomized to provide either BA or CT. Although a specified protocol detailing the allocation procedure to avoid confounding by therapist availability (e.g., more than one therapist available) was not employed, such confounding is highly unlikely. Thus, our trial approached quasi-randomization. Following assessment, patients were registered at the secured online research platform, which was used to collect the outcome measures at predefined time intervals.

### Interventions

We scheduled two therapy sessions per week during the first 8 weeks of the study. The remaining sessions were held on a weekly basis, the number of which depended on the progress of the individual patient. Although all therapists underwent formal CT training, most were relatively inexperienced. To optimize treatment and increase motivation to participate in this trial, they received additional training. This was provided in the form of separate two-day courses by two experienced clinicians in the field of CT and BA (S.D. Hollon and C. Martell). In addition, they received frequent peer supervision. Though a formal integrity check was not performed, treatment integrity was monitored by means of weekly peer consultation and occasional video-conferences with the two trainers. The CT approach was based on the original treatment manual developed by Beck and his colleagues (Beck et al. 1979) using adaptations

described elsewhere (Lemmens et al. 2011). Behavioral components were thus not specifically excluded in this approach. BA was provided according to the manuals developed by Martell and colleagues (Martell, Addis, and Jacobson 2001; Martell, Dimidjian, and Herman-Dunn 2010) as applied in a previous study on the effectiveness of BA (Moradveisi, Huibers, Renner, Arasteh, and Arntz 2013).

## Measurements

Participants completed all measures using the online platform and personal log-in codes. Following baseline assessment, measures were taken during the active treatment phase at 4, 8, 12, and 16 weeks. As described before, by week eight most participants had received 16 therapy sessions. Follow-up measures were taken at weeks 20 and 24. Dutch translations of the original instruments were used.

**Depressive Symptoms** Severity of depression was measured using the BDI-II, a 21-item tool with demonstrable reliability and construct validity for the original and Dutch version (Beck et al. 1996; Kuhner, Burger, Keller, and Hautzinger 2007; Nolen and Dingemans 2004). Values on the inventory range from 0 to 63 with higher scores indicating more severe depressive symptoms. More recent studies suggest a score below 12 to indicate at most minimal depression or the point of remission (Riedel et al. 2010).

**Negative Cognitions** Maladaptive negative beliefs were assessed using two instruments. The 17-item version of the Dysfunctional Attitude Scale (DAS-17; de Graaf, Roelofs, and Huibers 2009) and the revised version of the Automatic Thought Questionnaire (ATQ-R; Kendall, Howard, and Hays 1989; Raes and Hermans 2011). The ATQ-R was used to complement the DAS-17 as a measure of more superficial cognitions. Higher scores on both instruments refer to more negative cognitions.

**Behavioral Activation** The Behavioral Activation for Depression Scale (BADS; Kanter, Mulick, Busch, Berlin, and Martell 2007; Raes, Hoes, Van Gucht, Kanter, and Hermans 2010) consists of four factors, each of which comprises a subscale. The activation subscale (BADS<sub>a</sub>, 7 items) measures the degree of activation as perceived by the patient. Here, higher scores refer to more activation. The avoidance/rumination subscale (BADS<sub>ar</sub>, 8 items) assesses patients' ruminative and avoidant behaviors with higher scores relating to more of these instances. The school/work impairment and social impairment subscales (BADS<sub>ws</sub> and BADS<sub>s</sub>, each 5 items) examine problems in the respective areas and higher scores indicate more impairment.

**Therapeutic Alliance** The therapeutic alliance was examined with the 12-item form of the patient-rated version of the Working Alliance Inventory (WAI-S-C; Hatcher and Gillaspay 2006). The instrument incorporates three subscales, each consisting of four items. These refer to the perceived congruence between therapist and client on (a) the goals of therapy (WAI-S-C<sub>g</sub>), (b) the assumption that the tasks employed in therapy will aid in problem resolution (WAI-S-C<sub>i</sub>), and (c) the quality of the relationship within the therapeutic dyad (WAI-S-C<sub>b</sub>). Higher scores on any subscale indicate a better relationship between practitioner and patient.

## Statistical Analysis

Analyses were conducted at significance level  $\alpha = .05$  and based on the intention-to-treat (ITT) sample. Approaches that are insensitive to missing data (such as multilevel modeling) were not feasible in this study due to the size of our sample. Accordingly, initial attempts to use such analyses failed. In order to enable the use of other analyses, multiple regression imputation was used to replace incomplete information on the outcome and process measures. The investigation of missing data revealed no systematic pattern, confirming the validity of the application of multiple regression imputation. Random numbers were generated using the Mersenne Twister option and the results from a total of five imputations were aggregated to compose the final values. To preserve the explanatory power of the multiple assessment points, repeated measures analysis of variance (RM-ANOVA) were employed. BDI-II scores measured at all seven time points (i.e., including both follow-up assessments) were used as within-subjects variables to assess the effectiveness of the treatment modalities. Treatment condition was coded as  $-.5$  for CT and  $.5$  for BA and then added as the between-subjects variable (as suggested by Kraemer, Kiernan, Essex, and Kupfer 2008). Violations of sphericity were countered by using the Greenhouse-Geisser correction. Similarly, the change and temporal sequence of each individual process measure over the course of the treatment phase (baseline to week 16) was subjected to a RM-ANOVA with treatment condition as between-subjects factor (to assess differential process change in the two treatment conditions). Significant interactions between time and treatment condition were further investigated using simple contrasts for differences between baseline and each following assessment point per condition. Between- and within-group effect sizes (Cohen's  $d$ ) were computed for two time points per condition. The former were based on the mean scores at the end of the treatment phase (week 16) and last mean scores at follow-up (week 24). The latter were calculated from the difference between mean scores at baseline and week 16 and baseline and follow-up at week 24 (within the CT or BA group) and corrected for correlations between means. Next, exploratory analyses were conducted to investigate the predictive power of early change in process measures on later change in depression reduction. To do this, mean centered change scores from baseline to scores achieved at week 4 were computed for all process measures to comprise early change on these variables. Late change on depression was defined as the difference between BDI-II scores at week 4 and the last assessment during treatment at week 16 (before follow-up). Using multiple linear regressions, late change in depression was then regressed on early change in process measures, controlling for baseline BDI-II scores. To investigate differential effects of BA and CT, treatment condition and its interaction with change on mean centered process measures was included in the equations. Bonferroni corrections were used to counter inflation of type-I error rates (i.e.,  $\alpha_{\text{corrected}} .05/10 \text{ comparisons} = .005$ ). As outlined before, it is important to assess the direction of change and therefore look at the temporal sequencing of it (see also Kraemer, Stice, Kazdin, Offord, and Kupfer 2001; Kraemer et al. 2002). To provide a basic analysis of this, vice versa regression analyses were conducted. Early change in depression severity from baseline to week 4 was regressed on late change in the process variables, i.e., from week 4 to week 16. Case wise diagnostics with a two standard deviations ( $sd$ ) criterion were used to identify potential outliers and analyses were rerun after these were controlled for.

## Results

### Participants

In total, 44 patients were eligible for participation. Only one (allocated to BA) dropped out of the study before the start of the first session. Forty-three participants were therefore left for assessment (CT = 23, BA = 20). Their age ranged from 19 to 72 with a mean of 39 years ( $sd = 14.5$ ). Baseline scores on the BDI-II identified them as a highly depressed sample with mean scores of 30.2 and 32.1 in the CT and BA group respectively. Randomization proved successful in the equal distribution of demographic and baseline characteristics of process measures and depression scores. A slight majority of participants was female (51.2%) and most were of Dutch origin (95.3%). Table 1 provides an overview of demographic data per treatment condition. Inconsistent provision of information on depressive symptoms and processes was substantial, making imputation necessary. In the CT condition, 8.7% of data on all outcome and process measures were missing in week 4, further increasing to 34.8% in week 8. During weeks 12 to 20, the proportion sunk slightly to 30.4% only to increase at the last assessment in week 24 to 39.1%. In the BA condition, 5% of information on the ATQ-R-NL was missing at baseline assessment. For the remaining time points, the proportion was equal for all variables. From 20% in week 4, progressing to 35% in week 8 and slightly declining to 25% and 20% in weeks 12 and 16 respectively. At the first follow-up assessment, 35% of participants did not provide information on outcome measures, though only 30% of this information was missing at the second follow-up.  $\chi^2$  tests indicated no

**Table 1** Summary of demographic variables per condition

	CT ( $N = 23$ )	BA ( $N = 20$ )
Sex (% females)	52.2%	50%
Age ( $M, sd$ )	35.3 (12.6)	42.2 (15.9)
Relationship status ( $N, \%$ )		
Single	6 (26.1%)	5 (25%)
Living with family	4 (17.4%)	0 (0%)
Relationship/married	11 (47.8%)	12 (60%)
Divorced	2 (8.7%)	3 (15%)
Education ( $N, \%$ )		
None/elementary	2 (8.7%)	3 (15%)
Secondary	11 (47.8%)	5 (25%)
High-school	9 (39.2%)	11 (55%)
University	1 (4.3%)	1 (5%)
Employment ( $N, \%$ )		
Full-/part-time	10 (43.5%)	7 (35%)
Homemaker	1 (4.3%)	0 (0%)
Student	3 (13.1%)	3 (15%)
Unemployed/retired	9 (39.1%)	10 (50%)

differences in these percentages between the two conditions at any time point (all  $p \geq .28$ ).

At the end of the study, some patients were still in treatment. In these cases, the focus of therapy generally shifted towards issues not covered by the initial diagnosis of depression. Table 2 depicts a summary of descriptive statistics of depression and process measures after missing data were imputed. Based on the available data, the mean number of sessions received over the whole assessment period (i.e., including follow-up) was 21.3 ( $sd = 20.1$ ,  $N = 19$ ) in the CT condition and 19.6 ( $sd = 13.7$ ,  $N = 18$ ) in BA.

## Symptom Reduction

The results of the RM-ANOVA showed a significant reduction in depression scores over time,  $F(4.4, 179.8) = 39.7$ ,  $p < .001$ . However, no differences between conditions emerged,  $F(4.4, 179.8) = .96$ ,  $p = .44$ . The mean score on the BDI-II at the end of the study treatment phase in week 16 was 18.4 ( $sd = 11.6$ ) for CT and 20.2 ( $sd = 10.6$ ) for BA. At the last follow-up assessment in week 24, these scores were 15.4 ( $sd = 9.1$ ) for participants receiving CT and 14.4 ( $sd = 7.9$ ) for those in the BA group. The associated within-groups effect sizes at the last assessment were large with a Cohen's  $d$  of 1.89 for CT and 1.69 for BA (see Table 2 for all effect sizes). Figure 1a shows the rate of change on depression scores over the entire course of the study for both conditions.

## Change in Process Measures over Time

The results of the RM-ANOVA showed that all process measures underwent significant change over time. Change in the direction that indicates improvement was found for the DAS-17,  $F(4, 164) = 6.7$ ,  $p < .001$ ; the BADS total score,  $F(4, 164) = 11.6$ ,  $p < .001$ ; the BADS activation subscale,  $F(4, 164) = 6.0$ ,  $p < .001$ ; the WAI-S-C goal subscale,  $F(3.4, 139.1) = 3.5$ ,  $p = .01$ ; the WAI-S-C task subscale,  $F(4, 164) = 2.4$ ,  $p = .05$ ; the WAI-S-C bond subscale,  $F(3.2, 131.2) = 6.5$ ,  $p < .001$ ; and scores on the ATQ-R-NL,  $F(3.1, 128.5) = 12.1$ ,  $p < .001$ . Change in a negative direction (e.g., more ruminative behavior) was found on the BADS avoidance/rumination subscale,  $F(4, 164) = 7.3$ ,  $p < .001$ ; the BADS work/school impairment subscale,  $F(2.9, 118.7) = 7.6$ ,  $p < .001$ ; and the BADS social impairment subscale,  $F(3.1, 127.4) = 6.2$ ,  $p < .001$ .

## Interactions with Treatment

Two significant interactions emerged between treatment condition and time predicting scores on process measures. One was the DAS-17,  $F(4, 164) = 2.5$ ,  $p = .04$ : negative cognitions decreased in both conditions, but this change occurred earlier in BA compared to CT (see also Fig. 1b). This was confirmed by the simple contrasts comparing change on the DAS-17 from baseline to each subsequent assessment point. For the CT condition, the planned comparisons became significant at week 16 and remained so at both follow-up assessments (all  $p < .02$ ), whereas in the BA group, these contrasts were significant throughout the course

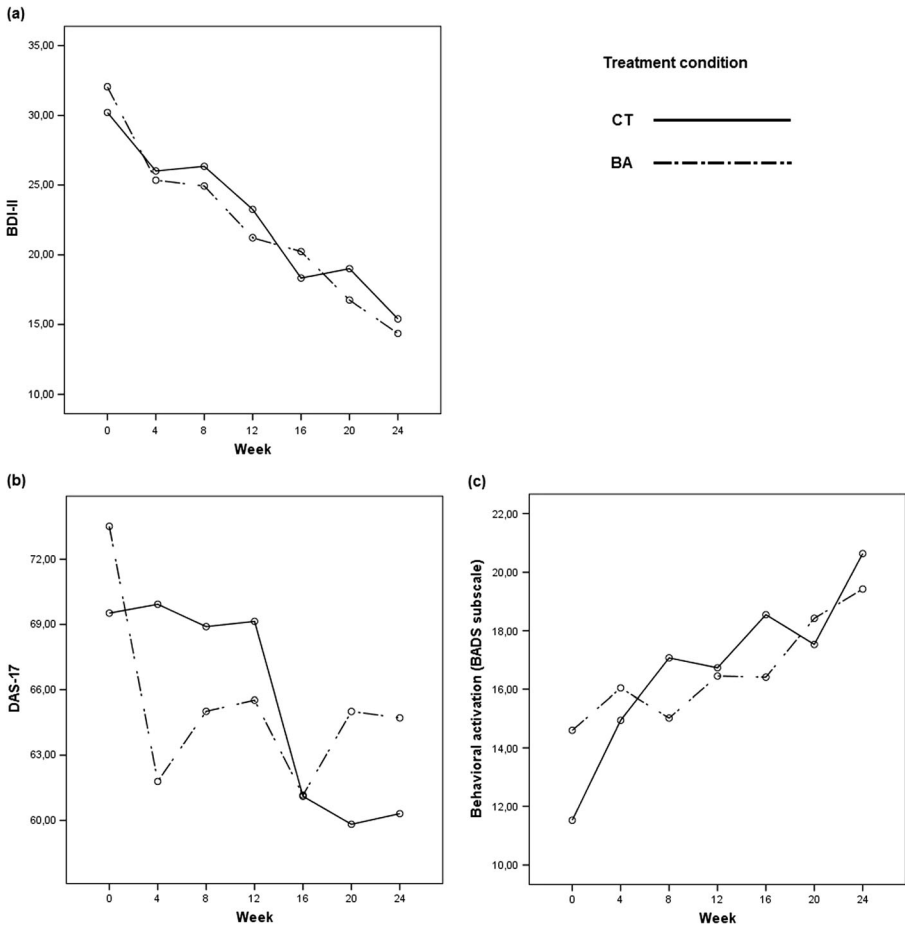
**Table 2** Descriptive statistics at baseline end of treatment, and follow-up, as well as within- and between-groups effect sizes for these time points

Variable and condition	Baseline		Week 16 (end of study treatment phase)				Week 24 (follow-up)			
	<i>M</i>	<i>sd</i>	<i>M</i>	<i>sd</i>	<i>d<sub>within</sub></i>	<i>d<sub>between</sub></i>	<i>M</i>	<i>sd</i>	<i>d<sub>within</sub></i>	<i>d<sub>between</sub></i>
BDI-II										
CT	30.2	8.1	18.3	11.6	1.5	-.17	15.4	9.1	1.89	.12
BA	32.1	11	20.2	10.6	1.04		14.4	7.9	1.69	
DAS-17										
CT	69.5	18.1	61.1	19.1	0.57	-.01	60.3	22.2	0.54	-.22
BA	73.5	18.6	61.2	17.4	0.76		64.7	16.2	0.45	
BADS <sub>t</sub>										
CT	74.9	16	96.1	15	-.95	.56	99.6	19.6	-1.0	.13
BA	72	18.1	85.7	21.7	-0.69		97.3	15	-1.38	
BADS <sub>a</sub>										
CT	11.5	4.8	18.5	5.9	-1.09	.34	20.6	6	-1.21	.2
BA	14.6	6.5	16.4	6.4	-0.28		19.4	5.7	-0.64	
BADS <sub>ar</sub>										
CT	26.5	8.1	32.9	6	-.59	.54	33.8	7.2	-.62	.28
BA	22.5	8.2	28.8	9	-.92		31.6	8.4	-1.26	
BADS <sub>ws</sub>										
CT	19	5	21.1	3.8	-.4	.23	22.4	4.3	-.61	-.24
BA	16.5	6.1	20	5.9	-.58		23.3	2.9	-1.6	
BADS <sub>s</sub>										
CT	17.8	6.2	23.5	3.1	-.87	.7	22.7	4.7	-.72	-.07
BA	18.5	5.5	20.7	4.8	-.32		23	3.1	-1.12	
WAI-S-C <sub>g</sub>										
CT	10.9	3	12	2.7	-.34	-.31	12.6	1.8	-.77	.14
BA	11.5	1.7	12.7	1.6	-.64		12.3	2.4	-.34	
WAI-S-C <sub>t</sub>										
CT	5.6	1.5	6.4	1.5	-.45	-.08	6.2	1.2	-.38	.00
BA	6.1	1.2	6.5	.9	-.29		6.2	.9	-.08	
WAI-S-C <sub>b</sub>										
CT	17	3.9	19	3.4	-.52	-.36	19.6	2.4	-.74	.45
BA	17.3	2.8	20	1.8	-1.17		18.6	2	-.41	
ATQ-R-NL										
CT	103.2	17.7	86.1	19.7	0.74	-.24	80.7	15.5	1.05	-.20
BA	112	22.3	91.4	24.3	1.22		83.8	14.7	1.59	

*N* = 23 (CT), 20 (BA)

of the study ( $p \leq .04$ ) and still marginally significant at the second follow-up in week 24 ( $p = .06$ ). The second significant interaction was found for the BADS activation subscale,  $F(4, 164) = 2.6$ ,  $p = 0.4$ : activation increased more steeply in CT, whereas this was relatively stable in BA (see also Fig. 1c). The simple contrasts





**Fig. 1** Change for scores on (a) the BDI-II, (b) the DAS-17, and (c) the BADS activation subscale for both conditions

for the BADS activation subscale showed that the planned comparisons from baseline to each subsequent assessment point were significant throughout all measurement points for CT (all  $p \leq .01$ ). In BA, the contrasts were only significant for the two follow-up measures (all  $p \leq .04$ ).

As shown in Table 2, within-group effect sizes of the process measures were generally moderate to large according to cut-points suggested by Cohen (1988). Merely, the effect of the WAI-S-C task subscale in week 24 almost disappeared in the BA condition, confirming the results of the analysis of change over time.

The effect sizes of all measures between the two groups were largely small (see Table 2). Taking into account the orientation of each instrument (i.e., do higher scores refer to improvement or impairment), a moderate effect was found on the BADS total score at week 16 ( $d = .56$ ), indicating more overall activation in the CT condition at that point, which is confirmed by the results on the activation subscale at the same time ( $d = .34$ ). The effect on the avoidance/rumination subscale ( $d = .54$ ) shows less of these instances and therefore more favorable outcomes for the BA condition at week 16, and

the same was true for the social impairment subscale ( $d = .7$ ). However, both of these were diminished or disappeared at follow-up. At week 24, one moderate effect emerged on the WAI-S-C bond scale ( $d = .45$ ), indicating a better relationship between patient and therapist in the CT condition. This effect reversed from the time of the last assessment during the study treatment phase in week 16 ( $d = -.36$ ).

### Prediction of Late Change in Depression

Early change on process measures was not predictive of late change in depression for any of the variables. This was true even before Bonferroni corrections were applied (all  $p > .36$ ). Removing two outliers (one from each treatment condition, identical cases for all measures) did not significantly change these results. Moreover, treatment condition and its interaction with early change on each process measure did not predict late change in depression (all  $p > .11$ ). Similar results were obtained when vice versa regressions were conducted in order to predict late change in process measures from early change in BDI-II scores (all  $p > .34$ ).

## Discussion

### Main Findings

We observed significant reductions over time on the BDI-II as a measure of depression in both conditions (although in the absence of differences relative to any control condition we cannot rule out the occurrence of spontaneous remission). At the end of the study, many participants still evidenced residual complaints and were still in treatment, which suggest that patients in routine settings need more and continued treatment to recover from depression. There was no statistically significant difference in BDI-II scores between the CT and BA group. Although this is in line with meta-analytic evidence (Barth et al. 2013; Cuijpers et al., 2013b), the twice-weekly session frequency could have led to an accelerated symptom reduction, distinguishing the effects from other studies that used different or similar frequency regimens. In this regard, the study presented herein served as a benchmark for the randomized controlled trial by Lemmens et al. (2015). We observed similar reductions in depression scores during a shorter period of time using twice-weekly sessions compared to the weekly sessions of either CT or IPT in the trial by Lemmens et al. (2015).

Despite the promising results for the early stages of treatment, they should be considered with caution. In their seminal randomized controlled trial, Dimidjian et al. (2006) compared the effects of CT, BA, and ADM for the treatment of depression. They found that for the highly depressed participants, BA outperformed CT. Moreover, on average, their participants experienced a reduction of around 13 points on the BDI-II in the CT and around 20 in the BA condition during the early therapy phase, therefore significantly exceeding the reductions found in our trial. A potential reason could be different levels of therapist experience and skill between the trials. Moreover, adherence to the treatment manual could not be monitored continuously since there was no formal integrity check.

## Process Measures

All processes that were hypothesized to be associated with symptom improvement changed over the course of treatment but mostly in a manner independent of treatment modality. It is conceivable that these processes were influenced directly by the relatively early changes in symptomatology and could therefore not be reliably distinguished as mechanisms of treatment.

Only negative cognitions measured by the DAS-17 and behavioral activation by the respective BADS subscale interacted significantly with treatment condition in the RM-ANOVA, pointing to a differential change path in BA and CT. However, as shown in Fig. 1b, the effect on the DAS-17 was largely a consequence of a sudden drop in maladaptive cognitions during the initial phase of BA. The post-hoc contrasts confirmed this relatively early reduction on negative cognitions in the BA group. This association is not what would have been predicted by cognitive theory. However, as Lorenzo-Luaces, German, and DeRubeis (2014) suggest in their extensive review, noncognitive interventions can produce changes in cognitions. This therefore does not disqualify cognitive change as a potential process underlying symptom reduction. In addition, depression is a complicated condition that involves a web of intertwined behavioral and cognitive factors. However, as is also shown in the figure, a comparable amount of change occurred on the DAS-17 in the CT condition at a later stage. Thus, the two treatment modalities were associated with comparable change in the DAS-17 over time, with the bulk of the observed change in BA happening over the first half of the study and the bulk of the change in CT happening over the second half. These findings are in line with previous research, showing that maladaptive thinking patterns “wax and wane” with the condition (Beevers and Miller 2005). Similarly, overall behavioral activation measured by the BADS activation subscale increased more in the CT group, as indicated by the RM ANOVA and effect sizes. This finding seems to counter the theoretical underpinnings of BA. Though the reasons for this are unknown, it is noteworthy that the CT approach did not specifically exclude behavioral techniques. In addition, cognitive restructuring techniques can lead to behavioral change. A general increase in activity was therefore not surprising.

An unexpected finding was the fact that most of the other constructs measured by the BADS changed significantly, but in a way that represents more problems in areas such as work, school, or other social contexts, and an increase in avoidant and ruminative behavior. Relatively more issues in the social domain were experienced by patients in the CT group. One explanation for this could be that they reported a larger increase in overall behavioral activation. More specifically, withdrawal from social situations has long been noted as a concurrent symptom in depression (Heinrich and Gullone 2006). Breaking through this containment is therefore certainly an uneasy process. In addition, healthy individuals often react rather negatively to depressed persons and their behavior (Sacco and Vaughan 2006), making the step out of the isolation even more difficult and confronting. The perception of an increase in problems in the social domain for patients in the CT group could therefore be explained by the fact that they encountered social situations more often. However, the between-groups effects on the BADS avoidance/rumination, work/school and social impairment subscales were either diminished or near zero at the follow-up assessment, showing that negative outcomes of such confrontation are not necessarily persistent.

The moderate effect size associated with the bond between therapist and patients suggests that at week 16, patients in the BA condition rate the quality of the relationship relatively better compared to those receiving CT but that this effect is reversed at follow-up. One possible interpretation is that the intensive consideration of thoughts and feedback by the therapist in CT increased the connection with the patient in the long term. On the other hand, the behavioral approach, using more applied exercises together with the therapist, could have enhanced the rating of the quality of the dyad particularly during the active treatment phase of the study, relative to the CT condition. Moreover, the bond between patient and therapist has been considered as a result of symptom improvement, rather than a mechanism in itself (Webb et al. 2011). The moderate size of the effect of this relationship at follow-up could therefore be a retrospection of some patients (as many of them did not receive treatment anymore at this point) that was the more positively evaluated, the better treatment outcomes were. However, as outlined before, others were still in therapy at this last assessment. For these patients, the WAI-S-C<sub>b</sub> was a more valid measure of the bond with the therapist.

### Strengths and Limitations

There were several limitations to this study that we tried to counter as far as this was possible. Although the design we used prevented us from drawing causal inferences, the finding that several of the processes that we observed (all derived from theory) did change in the predicted direction simultaneously with symptom change has an exploratory value. Moreover, the allocation procedure we used approximated quasi-randomization, an appropriate approach in this routine care setting. To counter usual criticisms of such pragmatic contexts, we specifically targeted and reduced the potential for the provision of lower quality therapy in effectiveness compared to efficacy trials. We did so by assuring good training and supervision by experts, as well as frequent peer meetings, although a formal integrity check was not performed. Despite these precautions, we cannot state that the level of therapists' experience was equivalent to the strict requirements normally applied in efficacy studies. Finally, the small sample size we recruited affects the inferences drawn. Attempts to implement more complex multilevel approaches for effectiveness and mediation analyses failed due to unsatisfactory model fit. However, the number of participants is a general issue in clinical research and our sample falls within a range comparable to some other studies (e.g., Troeung, Egan, and Gasson 2014; Weiss et al. 2012) and the chosen statistical approach is a sophisticated alternative in this context. Finally, missing data is often problematic and this trial suffered from a high degree of it. We used the best possible analytic approach in this situation, namely multiple imputation. Although analyses indicated no differences between the two conditions on the percentage of missing data, this does not mean that this is not a major concern. The results need to be interpreted with caution, as the imputation may have inflated results.

### Conclusion

Both CT and BA led to significant reductions in depressive symptomatology that was comparable to other studies and there is basic evidence that increasing the session frequency can have beneficial effects when compared to a benchmark trial. Moreover,

the comparable effectiveness of the BA and CT approaches adds to the literature, which finds BA to be potentially more easy to implement and more cost effective (Gilbody et al., 2017; Richards et al. 2016). However, there was no clear evidence of differential change with respect to purported underlying mechanisms. Although negative cognitions showed a time-lagged decrease in favor of BA, other analyses did not indicate differential change in the two conditions. The same was true for perceived activation (i.e., measured by the BADS activation subscale). That could have been the case if the respective theories are correct even in the absence of differences in outcomes between the treatment conditions. The overall finding that problems in the social domain as measured by the respective BADS subscales (i.e., avoidance/rumination, work/school, and social impairment) increased instead of decreased over the course of the study could be a result of overall activation.

To our knowledge, this is the first time a study on the comparison between both modalities employed multiple assessments of process measures in a pragmatic sample in routine practice. However, a limitation of this study is the relatively small sample we recruited, and we encourage other researchers to replicate our approach with a larger sample.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Barth, J., Munder, T., Gerger, H., Nuesch, E., Trelle, S., Znoj, H., et al. (2013). Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. *PLoS Medicine*, *10*(5), e1001454. <https://doi.org/10.1371/journal.pmed.1001454>.
- Beck, A. T. (1964). Thinking and depression: II. theory and therapy. *Archives of General Psychiatry*, *10*(6), 561–571. <https://doi.org/10.1001/archpsyc.1964.01720240015003>.
- Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: The Guilford Press.
- Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. (1996). Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *Journal of Personality Assessment*, *67*(3), 588–597. [https://doi.org/10.1207/s15327752jpa6703\\_13](https://doi.org/10.1207/s15327752jpa6703_13).
- Beevers, C. G., & Miller, I. W. (2005). Unlinking negative cognition and symptoms of depression: evidence of a specific treatment effect for cognitive therapy. *Journal of Consulting and Clinical Psychology*, *73*(1), 68–77. <https://doi.org/10.1037/0022-006x.73.1.68>.
- Brujniks, S. J., Bosmans, J., Peeters, F. P., Hollon, S. D., van Oppen, P., van den Boogaard, M., et al. (2015). Frequency and change mechanisms of psychotherapy among depressed patients: study protocol for a multicenter randomized trial comparing twice-weekly versus once-weekly sessions of CBT and IPT. *BMC Psychiatry*, *15*, 137. <https://doi.org/10.1186/s12888-015-0532-8>.
- Butler, A. C., & Beck, A. T. (1995). Cognitive therapy for depression. *The Clinical Psychologist*, *48*, 3–5.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: Lawrence Erlbaum Associates.
- Cuijpers, P., Andersson, G., Donker, T., & van Straten, A. (2011). Psychological treatment of depression: results of a series of meta-analyses. *Nordic Journal of Psychiatry*, *65*(6), 354–364. <https://doi.org/10.3109/08039488.2011.596570>.
- Cuijpers, P., Hollon, S. D., van Straten, A., Bockting, C., Berking, M., & Andersson, G. (2013a). Does cognitive behaviour therapy have an enduring effect that is superior to keeping patients on continuation pharmacotherapy? A meta-analysis. *BMJ Open*, *3*(4). <https://doi.org/10.1136/bmjopen-2012-002542>.

- Cuijpers, P., Sijbrandij, M., Koole, S. L., Andersson, G., Beekman, A. T., & Reynolds, C. F. (2013b). The efficacy of psychotherapy and pharmacotherapy in treating depressive and anxiety disorders: a meta-analysis of direct comparisons. *World Psychiatry, 12*(2), 137–148. <https://doi.org/10.1002/wps.20038>.
- Cuijpers, P., Smit, F., Oostenbrink, J., de Graaf, R., Ten Have, M., & Beekman, A. (2007). Economic costs of minor depression: a population-based study. *Acta Psychiatrica Scandinavica, 115*(3), 229–236. <https://doi.org/10.1111/j.1600-0447.2006.00851.x>.
- de Graaf, L. E., Roelofs, J., & Huibers, M. J. (2009). Measuring dysfunctional attitudes in the general population: the Dysfunctional Attitude Scale (form A) revised. *Cognitive Therapy and Research, 33*(4), 345–355. <https://doi.org/10.1007/s10608-009-9229-y>.
- DeRubeis, R. J., Evans, M. D., Hollon, S. D., Garvey, M. J., Grove, W. M., & Tuason, V. B. (1990). How does cognitive therapy work? Cognitive change and symptom change in cognitive therapy and pharmacotherapy for depression. *Journal of Consulting and Clinical Psychology, 58*(6), 862–869.
- DeRubeis, R. J., Gelfand, L. A., Tang, T. Z., & Simons, A. D. (1999). Medications versus cognitive behavior therapy for severely depressed outpatients: mega-analysis of four randomized comparisons. *The American Journal of Psychiatry, 156*(7), 1007–1013.
- DeRubeis, R. J., Hollon, S. D., Amsterdam, J. D., Shelton, R. C., Young, P. R., Salomon, R. M., et al. (2005). Cognitive therapy vs medications in the treatment of moderate to severe depression. *Archives of General Psychiatry, 62*(4), 409–416. <https://doi.org/10.1001/archpsyc.62.4.409>.
- Dimidjian, S., Hollon, S. D., Dobson, K. S., Schmalzing, K. B., Kohlenberg, R. J., Addis, M. E., et al. (2006). Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *Journal of Consulting and Clinical Psychology, 74*(4), 658–670. <https://doi.org/10.1037/0022-006x.74.4.658>.
- Fava, M., Bless, E., Otto, M. W., Pava, J. A., & Rosenbaum, J. F. (1994). Dysfunctional attitudes in major depression. Changes with pharmacotherapy. *The Journal of Nervous and Mental Disease, 182*(1), 45–49.
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002). Structured clinical interview for DSM-IV-TR axis I disorders, research version, patient edition. (SCID-I/P).
- Fried, E. I., & Nesse, R. M. (2014). The impact of individual depressive symptoms on impairment of psychosocial functioning. *PLoS ONE, 9*(2), e90311. <https://doi.org/10.1371/journal.pone.0090311>.
- Garratt, G., Ingram, R. E., Rand, K. L., & Sawalani, G. (2007). Cognitive processes in cognitive therapy: evaluation of the mechanisms of change in the treatment of depression. *Clinical Psychology: Science and Practice, 14*(3), 224–239. <https://doi.org/10.1111/j.1468-2850.2007.00081.x>.
- Gilbody, S., Lewis, H., Adamson, J., et al. (2017). Effect of collaborative care vs usual care on depressive symptoms in older adults with subthreshold depression: the casper randomized clinical trial. *JAMA, 317*(7), 728–737. <https://doi.org/10.1001/jama.2017.0130>.
- Greenberg, P. E., Fournier, A. A., Sisitsky, T., Pike, C. T., & Kessler, R. C. (2015). The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *Journal of Clinical Psychology, 76*(2), 155–162. <https://doi.org/10.4088/JCP.14m09298>.
- Hatcher, R. L., & Gillaspay, J. A. (2006). Development and validation of a revised short version of the working alliance inventory. *Psychotherapy Research, 16*(1), 12–25. <https://doi.org/10.1080/10503300500352500>.
- Heinrich, L. M., & Gullone, E. (2006). The clinical significance of loneliness: a literature review. *Clinical Psychology Review, 26*(6), 695–718. <https://doi.org/10.1016/j.cpr.2006.04.002>.
- Hofmann, S. G., Asnaani, A., Vonk, I. J. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: a review of meta-analyses. *Cognitive Therapy and Research, 36*(5), 427–440. <https://doi.org/10.1007/s10608-012-9476-1>.
- Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: a meta-analysis. *Journal of Counseling Psychology, 38*, 139–149. <https://doi.org/10.1037/0022-0167.38.2.139>.
- Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., et al. (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology, 64*(2), 295–304.
- Kanter, J., Mulick, P., Busch, A., Berlin, K., & Martell, C. (2007). The behavioral activation for depression scale (BADS): psychometric properties and factor structure.
- Kendall, P., Howard, B., & Hays, R. (1989). Self-referent speech and psychopathology: the balance of positive and negative thinking.
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 593–602. <https://doi.org/10.1001/archpsyc.62.6.593>.
- Kovacs, M., Rush, A. J., Beck, A. T., & Hollon, S. D. (1981). Depressed outpatients treated with cognitive therapy or pharmacotherapy. A one-year follow-up. *Archives of General Psychiatry, 38*(1), 33–39.

- Kraemer, H. C., Kiernan, M., Essex, M., & Kupfer, D. J. (2008). How and why criteria defining moderators and mediators differ between the Baron & Kenny and MacArthur approaches. *Health Psychology, 27*(2 Suppl.), S101–S108. [https://doi.org/10.1037/0278-6133.27.2\(Suppl.\).S101](https://doi.org/10.1037/0278-6133.27.2(Suppl.).S101).
- Kraemer, H. C., Stice, E., Kazdin, A., Offord, D., & Kupfer, D. (2001). How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors. *The American Journal of Psychiatry, 158*(6), 848–856. <https://doi.org/10.1176/appi.ajp.158.6.848>.
- Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry, 59*(10), 877–883.
- Kuhner, C., Burger, C., Keller, F., & Hautzinger, M. (2007). Reliability and validity of the Revised Beck Depression Inventory (BDI-II). Results from German samples. *Nervenarzt, 78*(6), 651–656. <https://doi.org/10.1007/s00115-006-2098-7>.
- Lemmens, L. H., Arntz, A., Peeters, F., Hollon, S. D., Roefs, A., & Huibers, M. J. (2015). Clinical effectiveness of cognitive therapy v. interpersonal psychotherapy for depression: results of a randomized controlled trial. *Psychological Medicine, 45*(10), 2095–2110. <https://doi.org/10.1017/s0033291715000033>.
- Lemmens, L. H., Arntz, A., Peeters, F. P., Hollon, S. D., Roefs, A., & Huibers, M. J. (2011). Effectiveness, relapse prevention and mechanisms of change of cognitive therapy vs. interpersonal therapy for depression: study protocol for a randomised controlled trial. *Trials, 12*, 150. <https://doi.org/10.1186/1745-6215-12-150>.
- Lewinsohn, P. M. (1974). A behavioral approach to depression. In R. J. Friedman & M. M. Katz (Eds.), *The psychology of depression: Contemporary theory and research* (pp. xvii–x318). Oxford: John Wiley & Sons.
- Lorenzo-Luaces, L., German, R. E., & DeRubeis, R. J. (2014). It's complicated: the relation between cognitive change procedures, cognitive change, and symptom change in cognitive therapy for depression. *Clinical Psychology Review, 34*(1), 1–10. <https://doi.org/10.1016/j.cpr.2014.12.003>.
- Martell, C. R., Addis, M. E., & Jacobson, N. S. (2001). *Depression in context: strategies for guided action*: Norton Press.
- Martell, C. R., Dimidjian, S., & Herman-Dunn, R. (2010). *Behavioral activation for depression: a clinician's guide*: The Guilford Press.
- Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of the therapeutic alliance with outcome and other variables: a meta-analytic review. *Journal of Consulting and Clinical Psychology, 68*(3), 438–450.
- Mathers, C., Fat, D. M., & Boerma, J. T. (2008). *The global burden of disease: 2004 update*: World Health Organization.
- Mehta, S., Mittal, P. K., & Swami, M. K. (2014). Psychosocial functioning in depressive patients: a comparative study between major depressive disorder and bipolar affective disorder. *Depression Research and Treatment, 2014*, 302741. <https://doi.org/10.1155/2014/302741>.
- Moradveisi, L., Huibers, M. J., Renner, F., Arasteh, M., & Arntz, A. (2013). Behavioural activation v. antidepressant medication for treating depression in Iran: randomised trial. *The British Journal of Psychiatry, 202*(3), 204–211. <https://doi.org/10.1192/bjp.bp.112.113696>.
- Nolen, W., & Dingemans, P. (2004). Meetinstrumenten bij stemmingsstoornissen. *Tijdschrift voor Psychiatrie, 46*(10), 681–686.
- Raes, F., & Hermans, D. (2011). De Nederlandstalige versie van de automatic thoughts questionnaire-revised (ATQ-R).[the Dutch version of the automatic thoughts questionnaire-revised (ATQ-R)]. *Gedragstherapie, 44*, 53–62.
- Raes, F., Hoes, D., Van Gucht, D., Kanter, J. W., & Hermans, D. (2010). The Dutch version of the behavioral activation for depression scale (BADs): psychometric properties and factor structure. *Journal of Behavior Therapy and Experimental Psychiatry, 41*(3), 246–250. <https://doi.org/10.1016/j.jbtep.2010.02.001>.
- Renner, F., Lobbestaël, J., Peeters, F., Arntz, A., & Huibers, M. (2012). Early maladaptive schemas in depressed patients: stability and relation with depressive symptoms over the course of treatment. *Journal of Affective Disorders, 136*(3), 581–590. <https://doi.org/10.1016/j.jad.2011.10.027>.
- Richards, D. A., Ekers, D., McMillan, D., Taylor, R. S., Byford, S., Warren, F. C., et al. (2016). Cost and outcome of behavioural activation versus cognitive behavioural therapy for depression (COBRA): a randomised, controlled, non-inferiority trial. *Lancet, 388*(10047), 871–880. [https://doi.org/10.1016/S0140-6736\(16\)31140-0](https://doi.org/10.1016/S0140-6736(16)31140-0).
- Riedel, M., Möller, H.-J., Obermeier, M., Schennach-Wolff, R., Bauer, M., Adli, M., et al. (2010). Response and remission criteria in major depression—a validation of current practice. *Journal of Psychiatric Research, 44*(15), 1063–1068. <https://doi.org/10.1016/j.jpsychires.2010.03.006>.
- Sacco, W. P., & Vaughan, C. A. (2006). Depression and the response of others: a social-cognitive interpersonal process model. *The interpersonal, cognitive, and social nature of depression, 101–132*.
- Spijker, J., Bockting, C., Meeuwissen, J., van Vliet, I., Emmelkamp, P., Hermens, M., & Van Balkom, A. (2013). Multidisciplinaire richtlijn Depressie (Derde revisie): Richtlijn voor de diagnostiek, behandeling en begeleiding van volwassen patiënten met een depressieve stoornis. [Multidisciplinary guidelines for

- depression (third revision): Guideline for the diagnosis, treatment and guidance of adult patients with a depressive disorder]. Utrecht: Trimbos Instituut.
- Teasdale, J. D. (1983). Negative thinking in depression: cause, effect, or reciprocal relationship? *Advances in Behaviour Research and Therapy*, 5(1), 3–25. [https://doi.org/10.1016/0146-6402\(83\)90013-9](https://doi.org/10.1016/0146-6402(83)90013-9).
- Trivedi, M. H. (2004). The link between depression and physical symptoms. *Primary Care Companion to The Journal of Clinical Psychiatry*, 6(suppl 1), 12–16.
- Troeng, L., Egan, S., & Gasson, N. (2014). A waitlist-controlled trial of group cognitive behavioural therapy for depression and anxiety in Parkinson's disease. *BMC Psychiatry*, 14(1), 19.
- Warmerdam, L., van Straten, A., Jongasma, J., Twisk, J., & Cuijpers, P. (2010). Online cognitive behavioral therapy and problem-solving therapy for depressive symptoms: exploring mechanisms of change. *Journal of Behavior Therapy and Experimental Psychiatry*, 41(1), 64–70. <https://doi.org/10.1016/j.jbtep.2009.10.003>.
- Webb, C. A., DeRubeis, R. J., Amsterdam, J. D., Shelton, R. C., Hollon, S. D., & Dimidjian, S. (2011). Two aspects of the therapeutic alliance: differential relations with depressive symptom change. *Journal of Consulting and Clinical Psychology*, 79(3), 279.
- Weiss, M., Murray, C., Wasdell, M., Greenfield, B., Giles, L., & Hechtman, L. (2012). A randomized controlled trial of CBT therapy for adults with ADHD with and without medication. *BMC Psychiatry*, 12(1), 30.

## Affiliations

**Felix Bolinski**<sup>1</sup> · **Gert-Jan Hendriks**<sup>2,3,4</sup> · **Seline Bardoel**<sup>5</sup> · **Steven D. Hollon**<sup>6</sup> · **Christopher Martell**<sup>7</sup> · **Marcus J. H. Huibers**<sup>1,8</sup>

<sup>1</sup> Department of Clinical, Neuro- And Developmental Psychology, Section of Clinical Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Van der Boechorstraat 7, 1081 BT Amsterdam, The Netherlands

<sup>2</sup> Center for Anxiety Disorders “Overwaal”, Institute for Integrated Mental Health Care Pro Persona, Nijmegen, The Netherlands

<sup>3</sup> Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands

<sup>4</sup> Department of Psychiatry, Radboud UMC, Nijmegen, The Netherlands

<sup>5</sup> Institute for Mental Health “Dr. Bosman”, Arnhem, The Netherlands

<sup>6</sup> Department of Psychology, Vanderbilt University, Nashville, TN, USA

<sup>7</sup> Department of Psychology, University of Massachusetts, Amherst, MA, USA

<sup>8</sup> Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA