

University of Massachusetts Medical School

eScholarship@UMMS

---

Preventive and Behavioral Medicine  
Publications

Population and Quantitative Health Sciences

---

2006-10

## A Qualitative And Quantitative Review of Behavioral Activation Treatment of Major Depressive Disorder

C. Richard Spates  
*Western Michigan University*

*Et al.*

Follow this and additional works at: [https://escholarship.umassmed.edu/prevbeh\\_pp](https://escholarship.umassmed.edu/prevbeh_pp)



Part of the [Behavioral Disciplines and Activities Commons](#), [Behavior and Behavior Mechanisms Commons](#), [Community Health and Preventive Medicine Commons](#), and the [Preventive Medicine Commons](#)

---

### Repository Citation

Spates CR, Pagoto SL, Kalata A.. (2006). A Qualitative And Quantitative Review of Behavioral Activation Treatment of Major Depressive Disorder. Preventive and Behavioral Medicine Publications. Retrieved from [https://escholarship.umassmed.edu/prevbeh\\_pp/113](https://escholarship.umassmed.edu/prevbeh_pp/113)

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in Preventive and Behavioral Medicine Publications by an authorized administrator of eScholarship@UMMS. For more information, please contact [Lisa.Palmer@umassmed.edu](mailto:Lisa.Palmer@umassmed.edu).

# A Qualitative And Quantitative Review of Behavioral Activation Treatment of Major Depressive Disorder

*C. Richard Spates, Ph.D., Sherry Pagoto, Ph.D., and Alyssa Kalata, B.A.*

In the treatment of Major Depressive Disorder (MDD), Behavioral Activation Therapy (BA) has emerged in recent years as an efficacious intervention. Derived from a component analysis of CBT, it offered at once a parsimonious explanation for the active ingredient of CBT, while demonstrating clinical efficacy as a separate treatment. Since the original investigation by Jacobson and colleagues in 1996, several well-controlled studies have been conducted, all of which converge to suggest strong support for BA as a stand-alone therapy for MDD. In this paper we review, evaluate and classify the evidence pertinent to this intervention and provide recommendations concerning its standing as a front line treatment. We conclude that the evidence is strong, the quality of research is generally very good and the effects sizes are compelling. Through both a qualitative and meta-analytic review of this evidence we also suggest the types of future studies that will establish greater confidence in BA as a front line treatment of choice therapy for Major Depressive Disorder.

Key Terms: Behavioral Activation, Major Depressive Disorder, Cognitive Behavior Therapy, Efficacious Treatment

---

Major Depressive Disorder (MDD) affects approximately 121 million people worldwide (World Health Organization, 2006). In the United States, MDD has a lifetime prevalence of 10% to 25% for women and 5% to 12% for men and a point prevalence of 5% to 9% for women and 2% to 3% for men (American Psychiatric Association, 2000). It has been determined to be the leading cause of disability and the 4<sup>th</sup> leading contributor to the global burden of disease (World Health Organization, 2006). Despite both the high prevalence and serious costs of depression, up to 90% of depressed patients in primary care settings fail to receive adequate treatment, in the form of empirically supported medication regimens or in the form of empirically supported psychotherapies for depression (Center for the Advancement of Health, 2000).

Of the available psychological therapies for treating MDD, Cognitive-Behavior Therapy (CBT) has emerged as the frontline treatment of choice, based on empirical findings demonstrating its efficacy and effectiveness. Dobson (1989) conducted a meta-analysis of the literature investigating Beck's Cognitive Therapy (CT) for depression (which is classified as Cognitive-Behavioral Therapy) and found that CT is more effective than general psychotherapy, Behavior Therapy, and pharmacotherapy in the treatment of MDD. Based in part on Dobson's conclusions, Chambless and colleagues (1998) labeled Cognitive Therapy (CT) for depression as a "well-established treatment." Recently, a component of CBT called Behavioral Activation (BA) has received a great deal of attention. First, because as a result of a component analysis it appeared to offer a conceptually parsimonious accounting of the measured effects of CBT. Secondly it offered potential as a more efficient and thus, cost-effective treatment for MDD. Both of these observations resulted from the work of Jacobson and colleagues (1996). For this reason the National Institute of Mental Health called for greater research attention to this technique (NIMH vision statement, 2005).

Although BA has a strong theoretical foundation (Lewinsohn, Sullivan, & Grosscup, 1980) and the packaged therapy from which it is derived (CBT) has strong empirical support, the studies devoted to the efficacy and effectiveness of BA as a stand-alone therapy, although encouraging, are somewhat limited. An extensive review of the PsychInfo database yielded eleven treatment relevant studies devoted

to behavioral activation, two of which were case studies (Hopko, Lejuez, & Hopko, 2004; Mulick & Naugle, 2004). One of the remaining investigations constituted a follow-up examination for an earlier treatment outcome study (Gortner, Gollan, Dobson, & Jacobson, 1998). The present paper provides a critical review of the controlled studies, and provides a meta-analytic summary of their collective findings. The purpose is to help frame policy recommendations at this point in time with respect to BA as a front-line stand-alone treatment.

We were guided in our review by the use of “gold standards” for treatment outcome studies outlined by Foa and Meadows (1997), and the system for classifying Level of Evidence used by the Agency for Health Care Policy and Research (AHCPR). Foa and Meadows identified seven criteria that characterize high quality treatment outcome research. These included:

1. Clearly defined target symptoms, which involves determining the diagnostic status of potential participants, specifying a threshold of symptom severity necessary for participation in the study, and determining inclusion and exclusion criteria.
2. Measures with appropriate levels of reliability and validity.
3. Blind evaluators should be employed.
4. Appropriately and adequately trained assessors.
5. Detailed treatment manuals should be used to guide conduct of each intervention.
6. Random assignment or stratified random assignment procedures to assign participants to treatment conditions.
7. Treatment adherence assessment to ensure that treatments were carried out as outlined in the treatment manual and to ensure that little or no treatment drift occurred.

The AHCPR classification system identifies 5 lettered ratings that characterize the level of evidence supporting health care interventions. These include:

1. Level A – Evidence is based on randomized, well-controlled clinical trials
2. Level B – Evidence is based on well-designed clinical studies, without randomization or placebo comparison.
3. Level C – Evidence is based on service and naturalistic clinical studies, combined with clinical observations that are sufficiently compelling to warrant use of the treatment technique or follow the specific recommendation.
4. Level D – Evidence is based on longstanding widespread clinical practice by circumscribed group of clinicians that has not been subjected to empirical tests.
5. Level F – Evidence is based on recently developed treatment that has not been subjected to clinical or empirical tests.

We provide a description of the studies in sufficient detail to permit a “gold standard” analysis, followed by a meta-analytic review of the quantitative evidence. We also apply an AHCPR classification level for each study. We then provide an overall summary of the state of the evidence at this time and recommend future research designed to strengthen confidence in this intervention as a front-line intervention for MDD.

Jacobson, Dobson, Truax, Addis, Koerner, Gollan, and colleagues (1996) investigated the effectiveness of BA in the context of a component analysis of CT. 150 outpatients diagnosed with MDD were randomly assigned to receive one of three treatments, all of which were components of Beck’s CT package: BA, BA with a skill teaching component designed to modify automatic thoughts (AT), or the full CT package. Assessments were conducted by trained and supervised graduate students prior to therapy, at the time of termination of therapy, and at 6, 12, 18, and 24-month follow-up examinations. Participants were randomly assigned to one of the three aforementioned treatment groups based upon

clear inclusion criteria. The BA treatment condition involved daily activity monitoring, assessment of pleasure and mastery when engaging in various activities, assignment of tasks of increasing difficulty, cognitive rehearsal and discussion of scheduled activities, discussion of specific problems and provision of behavior therapy techniques to remedy the problems, and social skills training. The AT treatment condition involved all of the components from the BA condition, in addition to noticing mood shifts during session and inquiring about the thoughts that occurred before the mood shift, completing daily thought records, reexamining and discussing thoughts clients made during certain situations and determining if the conclusions made were warranted, helping clients to learn to respond more appropriately to negative thinking, and examining attributional mistakes in terms of successes and failures in the patients' lives. The CT condition involved all the components from the AT condition, in addition to use of the "downward arrow" technique, the identification of core beliefs, the identification of alternative core beliefs, and discussion of the advantages and disadvantages and short-term and long-term consequences of certain core beliefs. All treatments were administered by four experienced cognitive therapists, all of whom had participated in at least one previous clinical trial in which they served as cognitive therapists and all of whom had been trained for a year prior to the beginning of the study. The data obtained from the study indicated that CT was no more effective than BA or AT. There were no significant differences between the two primary outcome measures at post-test or six-month follow-up. All groups had maintained most of the gains made in treatment at six-month follow-up. There were also no significant differences between the number of individuals meeting criteria for improvement or recovery. Finally, the BA and AT conditions were equally as effective as the CT condition at modifying negative thoughts and maladaptive attributional styles, though in the case of BA these features were not specifically targeted.

Gortner, Gollan, Dobson, and Jacobson (1998) presented the two-year follow-up data for the Jacobson et al. (1996) study. All three treatments led to lasting change, with no one treatment outperforming the other two. There was no significant difference between relapse rates or time to first relapse after treatment between the three groups. Finally, there was no significant difference between the number of "well weeks" (weeks during which participants experienced minimal or no depressive symptoms), with 75% of the weeks during the two-year follow up categorized as "well weeks" across all three groups.

The Jacobson et al. (1996) investigation met all seven of the "gold standards" of treatment outcome studies. Diagnostic status was ascertained through the administration of the Structured Clinical Interview for DSM-III-R Axis I Disorders and severity was determined through scores on the Beck Depression Inventory and the Hamilton Rating Scale for Depression. Inclusion and exclusion criteria were clearly delineated. The two primary outcome measures for the study, the Beck Depression Inventory (BDI) and the Hamilton Rating Scale for Depression (HRSD), have good psychometric properties. The individuals conducting assessments were extensively trained and supervised and were not the same individuals that were administering treatment. Furthermore, interrater reliability was .9 for diagnoses made using the Structured Clinical Interview for DSM-III-R Axis I Disorders (SCID-IV). Random assignment procedures were employed to assign individuals to treatment groups. Manuals were completed for each of the treatment groups and were based on the original CT manual. Finally, therapists were extensively monitored for adherence to treatment protocol. Treatment integrity was assessed by independent raters who listened to audiotapes of 20% of the treatment sessions. Any protocol violations were expeditiously addressed. Monthly meetings also occurred with two of the authors to discuss protocol and treatment integrity issues. In addition to monitoring for adherence, therapist competence was also measured using the Cognitive Therapy Scale, and all four therapists were above cutoff scores necessary to determine competence. This investigation warrants an AHCPR rating of "A", indicating that the "evidence is based upon a randomized, well-controlled clinical trial for individuals" with MDD.

Hopko, Lejuez, LePage, Hopko, and McNeil (2003) conducted a randomized pilot trial of Brief BA with 25 depressed individuals in an inpatient psychiatric hospital. Assessments using the BDI and diagnostic interviews were conducted prior to the first session. The BDI was repeated at post-treatment on day 14 or at discharge. Patients were randomly assigned to receive Brief BA or standard supportive treatment typically provided in the hospital. In both conditions, patients were enrolled in treatment for two weeks or until discharge and during this time, were seen three times per week for 20 minutes by a master-level clinician. The master level clinicians had extensive training and experience with cognitive-behavioral interventions and were monitored for adherence to the treatment protocol through weekly supervision meetings with a licensed clinical psychologist with familiarity with BA principles. All patients also received antidepressant medication. In the Brief BA group the function of depressed behavior was assessed, strategies for reducing reinforcement for depressed behavior were introduced, a treatment rationale was provided, and an activity hierarchy was constructed and then followed in a systematic fashion. Furthermore, tokens were provided to patients contingently on the basis of completion of Brief BA for treatment-related goals. In the standard supportive treatment group, patients engaged in nondirective discussion in which they were encouraged to share their experiences and discuss their problems by their clinician. Patients in this condition were yoked to individuals in Brief BA such that they received tokens in a noncontingent manner. The data collected indicated that individuals in the Brief BA group had significantly greater improvement in BDI scores from pre-treatment to post-treatment than those individuals in the standard supportive treatment condition. Furthermore, a large effect size was obtained for individuals in the Brief BA group.

The Hopko et al. (2003) study clearly met two, and possibly three, of the seven “gold standards” for treatment outcome studies. The only outcome measure used in this study, the BDI, has adequate psychometric properties. Random assignment procedures were employed. In terms of treatment integrity, the individuals administering the behavioral activation intervention had weekly supervision meetings with a licensed psychologist with extensive knowledge of the behavioral activation protocol. The expectation was that these supervision meetings would be used to monitor treatment adherence. However, no *measures* of treatment integrity were reported. . This investigation warrants an AHCPR rating of “A/B”, indicating that the Evidence is based on a well-designed clinical study, without randomization or placebo comparison.” The study did not include a complete description of the inpatient sample and lack integrity measures that would have been expected of a higher classified trial.

Cullen, Spates, Pagoto, and Doran (2006) randomly assigned 25 individuals diagnosed with MDD (using the Structure Clinical Interview for DSM-IV or SCID-IV) to 10 weeks of individual BA therapy or 6 weeks of no treatment in the waitlist-control condition. Assessments, using the SCID-IV, BDI and Revised Hamilton Rating Scale for Depression (RHRSD) were conducted at pre-treatment, post-treatment, and three-month follow-up by doctoral students in clinical or counseling psychology. Furthermore, individuals who were assigned to the waitlist-control condition were assessed bi-weekly with the BDI – II and also given a full assessment (BDI, RHRSD) at the end of the wait period. Individuals once assigned to the BA treatment group, were administered the BDI – II on a weekly basis. The BA treatment protocol used was similar to that implemented by Jacobson and colleagues (1996). Sessions involved the administration of the BDI– II, a review of homework, addressing topics on the agenda for the day, a review of the topics covered, and the assignment of homework. Those individuals assigned to the waitlist condition were told that they were undergoing an “assessment phase” prior to treatment. These individuals came to the clinic on a bi-weekly basis for six weeks in order to complete the BDI– II. The conclusion were that BA led to clinically and statistically significant reductions in levels of depression symptoms (BDI-II and HRSD) and formal diagnoses of MDD (SCID-IV) in comparison to the waitlist comparison.

The Cullen et al. (2006) study meets all seven of the “gold standards” of treatment outcome studies. The SCID-IV was used to assess if participants met criteria for MDD and the BDI – II, and the



RHRSD were used to assess the severity of depression symptoms. Inclusion and exclusion criteria were explicitly described. Random assignment procedures were employed. The three primary outcome measures used in this study all have adequate psychometric properties. Evaluators were not the same individuals who were administering treatment and were adequately trained. Therapists were trained for 12 hours in behavioral activation techniques and were then given protocol outlines to guide the administration of treatment. Treatment adherence was monitored through videotape review of 11% of all therapy sessions. Review of these tapes indicated that therapists were adhering strongly to treatment protocol. In addition, weekly research team meetings were held to discuss individual cases and as a means of preventing therapist drift. This investigation warrants an AHCPR rating of “A”, indicating that the “evidence is based upon a randomized, well-controlled clinical trial for individuals” with MDD. However the study suffered from a small sample size and warrants replication to establish confidence in its conclusions.

#### *Studies of group BA*

Porter, Spates, and Smitham (2004) investigated the effectiveness of BA group therapy (BAGT) for individuals seeking mental health services for depression at rural community mental health agencies. Assessments using the SCID-IV, RHRSD and the BDI-II were conducted at pre-treatment, post-treatment, and three-month follow-up. Additionally, the BDI- II was administered on a weekly basis for all individuals once assigned to BAGT. Individuals who were assigned to the waitlist group were also assessed at the beginning and the end of the waiting period. 37 individuals were randomly assigned to either BAGT or a four- to six-week waitlist group. The BAGT condition was administered by eight staff therapists who had degrees in either social work or psychology and who were licensed professionals. The BAGT protocol was heavily based on that in the Jacobson et al. (1996) study. Treatment sessions were led by two co-therapists and attended by six to ten individuals. Sessions were 95 minutes long. Data from this study indicated that BAGT produced statistically and clinically significant reductions in symptoms of depression. Furthermore, those individuals assigned to the waitlist group did not improve during their time on the waitlist, but did improve once they received the BAGT. The data also indicated that there was a statistically significant decrease in the number of individuals who met diagnostic criteria for depression from pre-test to the three-month follow-up assessment.

The Porter, Spates, and Smitham (2004) study met six of the seven of the “gold standards” of treatment outcome studies. The SCID-IV was used to determine the presence or absence of a diagnosis of MDD. The BDI – II and the RHRSD were used to assess the level of symptom severity. Random assignment procedures were employed. Inclusion and exclusion criteria were clearly delineated. The measures used, all have adequate psychometric properties. Independent evaluators performed the assessments and were trained. A detailed BAGT manual gave step-by-step instructions on how to implement each phase of treatment and was provided to all therapists, along with booster sessions at 2-3 month intervals. Finally, although therapists were extensively trained, and the authors communicated regularly with therapists about delivering the intervention, no measures of treatment integrity were taken. This investigation warrants an AHCPR rating of “A”, indicating that the “evidence is based upon a randomized, well-controlled clinical trial for individuals” with MDD.

#### *Unpublished Studies of BA*

Wright, Spates, Bloem, DeViva, and Pope (2003) conducted a study assessing the effectiveness of behavioral activation administered in a group format to 62 veterans on a Specialized Inpatient PTSD Unit in a Veteran Affairs Medical Center. The participants presented with co-morbid PTSD and MDD. While staying on the Unit, patients typically completed three treatment tracks. The C-track was a 22-day program involving education and stabilization. The R-track was a 28-day program consisting of further education and beginning to address the trauma recovery process. The S-track was a 56-day program,

focusing specifically on trauma, and involving continued education. Although a therapy group met multiple times a week in each track, depression was not characteristically or explicitly addressed as part of treatment. In this study, the BAGT intervention was integrated in to the R-track. Prior to admission to the unit, all individuals were assessed to ensure that they met criteria for a diagnosis of PTSD. Assessments pertaining to the study were conducted at pre-treatment, post-treatment, one-month follow-up, and two-month follow-up. Upon entering the unit in a group cohort, participants were randomly assigned with the rest of their cohort to either the BAGT group or a comparison group, which received standard treatment of classes and group therapy. Three cohorts were assigned to the BAGT group and four cohorts were assigned to the comparison (Treatment as Usual) group. The BAGT group met twice a week and was run by a psychology intern and staff psychologist. In regard to content, two to three psychoeducational topics were discussed each week, including standard skill-sets indicated in the BA protocol. During the group therapy, comprehension of these topics was assessed, relevant examples were elicited, and homework was assigned and reviewed. In addition to receiving BAGT, individuals assigned to this condition also attended the same classes and group therapy offered to the comparison group. The results indicated that on all measures of depression, individuals in the BAGT treatment group showed significant reductions in symptoms of depression from pre-treatment to two-month follow-up, whereas individuals in the comparison group did not. In the BAGT group, there was an increase in scores on measures of depression from post-treatment to one-month follow-up, although these scores decreased by two-month follow-up. Neither group showed significant changes in measure of PTSD. Finally, this study found a strong positive correlation between patient completion of homework assignments and reduced symptoms at posttest and follow-up evaluations.

The Wright et al. (2003) investigation met four of the seven “gold standards” of treatment outcome studies. Although diagnostic status of PTSD and inclusion and exclusion criteria was assessed during the admissions process (instead of by the research team), a diagnosis of MDD and level of symptom severity were determined by the research team. The primary outcome measures used in the study (the BDI – II, the RHRSD, Posttraumatic Stress Diagnostic Scale (PDS), and the Mississippi Scale for Combat-Related PTSD (MSC-RPTSD) all have adequate psychometric properties. Assessors were properly trained, however because of their involvement on the unit, they may not have been blind to treatment condition. The manual used in this study was based on the manual used in the Jacobson et al. (1996), but was modified such that it was appropriate for use in group, rather than individual, therapy. Although cohorts of individuals were randomly assigned to treatment conditions, the individuals themselves were not. Finally, although therapists provided ratings of their own treatment adherence and although these ratings were relatively high (91% and 85%), no measures of treatment adherence were taken by independent evaluators. This investigation warrants an AHCPR rating of “A/B”, indicating that the “evidence is based upon a randomized, well-controlled clinical trial for individuals” with MDD. This investigation was limited by a relatively small sample size, integrity assessment, and a weaker randomization procedure, and thus replication is required to establish confidence in its findings.

Dimidjian (2005) conducted a study comparing the efficacy of BA, CT, antidepressant medication (Paroxetine), and pill-placebo in the treatment of MDD. Assessments were conducted at pre-treatment, mid-treatment (eight weeks), and post-treatment (16 weeks), and at other points during treatment as clinical indicated (for example, in instances of early termination). 241 individuals were randomly assigned to one of the four aforementioned conditions. Three therapists provided treatment in the BA condition. Therapists in this condition received both on-site and off-site supervision. Individuals assigned to this condition received a maximum of 24 50-minute sessions, with sessions held biweekly for the first eight weeks and weekly for the final eight weeks. The BA treatment was an expanded version of the BA treatment used in the Jacobson et al. (1996) investigation. The components borrowed from the treatment used in Jacobson et al. included self-monitoring, activity scheduling, ratings of pleasure and mastery when engaging in certain activities, exploring alternative behaviors related to goal-setting, and role playing to address behavioral deficits. The components added to treatment for this study included the

assessment and treatment of avoidance behaviors, establishing and maintaining a regular routine, the explicit targeting of rumination, a behavioral case conceptualization, and behavioral treatment rationale. Three therapists provided CT. All therapists in this condition had served as cognitive therapists in prior studies and were supervised both on-site and off-site on a weekly basis. Individuals in this condition received the same number and length of therapy sessions as those individuals assigned to the BA treatment. The CT provided in this study was the standard form recommended by Beck and as such, explicitly targeted behavioral dysfunction, situation specific negative thinking and cognitive distortions, and underlying dysfunctional beliefs. Five board-certified pharmacotherapists provided services in the antidepressant medication and pill-placebo conditions. These individuals were trained and supervised by an experienced pharmacotherapy researcher for the duration of the study. The two conditions were triple-blind for the first eight weeks of the study, but at eight weeks the blind was broken and individuals assigned to the pill-placebo condition were offered their choice of treatment. The antidepressant medication condition was then conducted in a single-blind fashion for the remaining eight weeks of the study. Individuals in these two conditions were seen weekly for the first four weeks and biweekly for the remainder of the study following the first four weeks. The first session was 30 to 45 minutes and the subsequent sessions were 30 minutes. Typical sessions included the administration of the RHRSD, inquiries about treatment response, side effects, non-study medications, and involved renewals or modifications of pill dosages. The findings from this study indicated that for those individuals who were in the high severity subgroup, BA and antidepressant medications performed equally well and outperformed CT. However, for individuals who were in the low severity subgroup, all treatments led to significant improvement over time, with no differential effects seen for any active treatment in particular.

The Dimidjian (2005) investigation met all seven of the “gold standards” for treatment outcome studies. A diagnosis of MDD was made through the use of the SCID-IV, and severity of depression was measured through the use of both the BDI- II and the RHRSD. . Stratified random assignment procedures utilizing symptom severity as the basis for stratification were employed to assign participants to groups. Inclusion and exclusion criteria were also explicitly stated. The two primary outcome measures in the study (the BDI- II and the RHRSD) both have adequate psychometric properties. All evaluators were independent of treatment administration, and all evaluators were blind to the treatment conditions. Furthermore, all evaluators were trained, certified, and monitored through weekly supervision meetings to prevent any evaluator drift. A detailed treatment manual was used. Finally, individuals who were blind to treatment condition monitored treatment adherence through ratings of the presence or absence of techniques appropriate to condition. Ratings of competence were also made in the CT and antidepressant medication conditions. Ratings indicated that therapists were adhering strongly to treatment protocol and were administering their respective treatments in a competent manner. This investigation warrants an AHCPR rating of “A”, indicating that the “evidence is based upon a randomized, well-controlled clinical trial for individuals” with MDD.

#### *Non-experimental studies*

Hopko, Armento, Hunt, Bell, and Lejuez (2005) conducted a preliminary clinical trial of Brief BA Treatment for Depression with six individuals diagnosed with cancer and MDD. Undergraduate research assistants and graduate students who were supervised by the principal investigator completed assessments at pre-treatment, post-treatment, and three-month follow-up. Supervision was conducted through the use of audiotapes and discussion, after which a consensus diagnosis was made. The Brief BA Treatment for Depression consisted of nine one-hour sessions covering psychoeducation, explanation of the treatment rationale, activity and goal selection, and behavioral activation. Treatment was conducted by two clinical psychology graduate students who were supervised by the principal investigator through an audiotape review process. Ratings of therapist competence and therapist adherence to the treatment protocol were conducted for 20% of the sessions and indicated that the therapists involved in the study administered the treatment competently and as outlined in the treatment protocol. The six patients



involved in the study received a treatment adherence score of 78%, indicating that the patients completed an average of 78% of the behavioral assignments. The data collected indicated significant pre-treatment to post-treatment change on measures of depression, quality of life, and medical outcomes, yielding moderate-to-large effect sizes. Furthermore, these gains were maintained at the three-month follow-up assessment, and the non-significant gains in somatic anxiety and bodily pain measures observed from pre-treatment to post-treatment assessment improved to clinically significant gains by this follow-up assessment.

The Hopko et al. (2005) investigation met five of the seven of the “gold standards” of treatment outcome studies. Diagnostic status was ascertained through the use of the SCID –IV, and a severity criterion for MDD was specified for inclusion in the study. Inclusion and exclusion criteria were clearly delineated. The primary outcome measures, which included the BDI – II, the RHRSD, and the Center for Epidemiological Studies of Depression Scale all have solid psychometric properties. The individuals who administered assessments were independent of treatment administration. No explicit mention of a treatment manual for use by the therapists was made by the authors. However, treatment integrity and therapist competence were closely monitored. The only “gold standard” criterion this study failed to meet was random assignment to groups. This study only had one group and as such, random assignment was not possible. This investigation warrants an AHCPR rating of “**B**”, indicating that the Evidence is based on well-designed clinical studies, without randomization or placebo comparison” with MDD.

#### *Summary of well-controlled trials*

Taken as a whole, this body of research provides good preliminary evidence for the effectiveness of BA as a treatment for depression. On average, these studies met approximately five of the seven of the “gold standards” for treatment outcome research, indicating that the quality of research on the efficacy and/or effectiveness of BA is adequate but that there is room for improvement. This small body of research had the additional constraint of small sample sizes in several instances, suggesting that more research, especially with larger sample sizes, is needed before BA can confidently be raised to the level of a frontline intervention of choice for MDD. Overall the line of investigation supporting this intervention warrants an AHCPR mixed classification of “A/B”. Three investigations achieved the highest gold standard ratings (7) and AHCPR classification (A), while most others achieve an AHCPR classification of “B”. The most consistent weakness across studies entailed small sample sizes. To that end, we now address the collective evidence vis-à-vis a meta-analytic review of these same investigations. This analysis will permit an examination of effect sizes for this novel intervention.

### Method

#### *Procedure*

Studies were identified via searches of Medline and PsychInfo from 1997 to 2006 using “behavioral activation” and “behavior therapy and depression” as keywords. Reference lists of studies pertaining to BA were also scanned for additional studies. Both published and unpublished studies were included in the analysis. Only treatment studies that included a baseline and at least one follow-up assessment point were included in the analysis. Most studies had end-of-treatment and 3-month assessment points, thus our analysis targeted change from baseline to each of these time points.

Eight studies were included. Only studies using a standardized measure of depression were included and all studies found met this criteria. The primary focus of this article is the within-subjects effects of BA for Major Depressive Disorder. Two studies randomly assigned participants to BA or wait-list control (Porter et al. 2004, Cullen et al., 2006) and four studies randomly assigned participants to BA or other treatments such as supportive psychotherapy (Hopko et al., 2003), usual care (Wright et al.,

unpublished), cognitive therapy (Jacobsen et al., 1996; Dimidjian, unpublished), or medication (Dimidjian, unpublished). The 2 remaining studies did not utilize random assignment and were single group designs (Hopko et al., 2005; Pagoto et al., unpublished).

### *Measures*

The primary outcome for all studies was depressive symptomatology which was most often measured using the BDI-II and the RHDS. Because all but one study utilized both of these measures and they are standardized measures of depression, the present analysis reports effect sizes from these measures.

### *Effect Size Calculation*

Effect sizes were calculated by computing pre- to post-treatment and pre-treatment to follow-up (3-months, unless otherwise stated) differences for the treatment group using Cohen's  $d$   $[(\bar{X}_{pre} - \bar{X}_{post}) / SD_{pooled}]$  and  $[(\bar{X}_{pre} - \bar{X}_{follow-up}) / SD_{pooled}]$ . Cohen's  $d$  was selected as the effect size estimate because it permits pre-test – post-test comparisons within studies, allowing for comparisons of effect sizes across studies that utilize different types of control conditions. Cohen's  $d$  tends to provide a more conservative estimate than other commonly used methods (Rosenthal, 1991). The small number of studies did not allow for BA versus control or BA versus other treatment comparisons.

Average effect sizes were derived within studies by computing the mean effect size for the Beck Depression Inventory and Hamilton Depression Rating Scale scores. Effect sizes were weighted by the sample size on which it was based when computing the mean of all effect sizes. Two studies (Hopko et al., 2003; Dimidjian, unpublished) only reported post-treatment scores and were therefore not included in the mean effect size calculation for follow-up time points. One study reported data on both low and high symptom severity levels (Dimidjian, unpublished), thus effect sizes for both subgroups are reported.

## Results of Meta Analysis

Effects sizes for improvement in BDI-II scores ranged from medium to very large (.49 to 4.03) at post-treatment and small to very large (.3 to 2.92) at follow-up (see Table 2). Of the 8 studies, 6 reported very large effect sizes ( $> 1.0$ ) for BDI-II scores at post-treatment and 5 of 6 reported very large effect sizes at follow-up. For HDRS scores effect sizes ranged from large to very large (.83 to 3.29) at post-treatment and large to very large (.96 to 2.48) at follow-up. All 7 studies utilizing the HDRS reported large or very large effect sizes at both post-treatment and follow-up.

### *Attrition*

Attrition rates for the 7 studies that involved outpatient treatment ranged from 9-33%, with smaller studies having higher rates of attrition, generally. The mean attrition rate across these 7 studies was 18.71%, which is comparable to attrition rates reported in other depression psychotherapy (Last 1985; Satterfield, 1998:) and pharmacotherapy research (Aikens, Nease, Nau, Klinkman, & Schwenk, 2005).

### *Summary of Meta Analysis*

To date, only a small number of treatment studies have been conducted on BA treatment for depression. Those studies have revealed significant and fairly large effect sizes on standardized measures of depression. In fact there appears to be a strong association of number of sessions required in the

protocol and effect size achieved (range 6 to 24 visits). The average attrition rate was fairly comparable with other psychotherapies and better than pharmacotherapies for depression.

Although these initial studies of the efficacy of BA have had consistently positive outcomes, larger randomized trials comparing BA to other therapeutic modalities are needed.

Table 1. Mean Pre- Versus Post-Treatment and Pre- Versus Follow-Up Effect Sizes

	Effect Size (d)		
	Weighted MeanSD		95% CI
<b>BDI-II</b>			
Pre-post	2.11	1.14	1.13, 3.15
Pre-Follow-up	1.83	.93	.78, 2.74
<b>HDRS</b>			
Pre-post	1.91	.88	1.23, 2.70
Pre-Follow-up	1.78	.64	1.06, 2.40

Table 2. Characteristics of studies included in the present meta analysis

Authors	Sample size (# receiving BA)	Comparison groups	Treatment modality	Number of visits	Assessment points	Measures	Effect size (pre-post, pre-follow-up)	Attrition rate
Cullen et al 2006	N=25 (25)	Wait-list control	Individual	10	Pre-, Post-treatment, 3 months	BDI-II, HDRS	BDI (1.82, 1.74) HDRS (.83, .96)	32%
Pagoto et al unpub'ed	N=14 (14)	None	Individual	12	Pre-, Post-treatment, 4-months	BDI-II and HDRS	BDI (3.19, 2.01) HDRS (2.82, 2.05)	14%
Dimidjian et al	N=241 (43)	Cognitive therapy, paroxetine, placebo	Individual	24	Pre-, post-treatment	BDI-II and HDRS	BDI low severity (2.41) high severity (4.03) HDRS low severity (2.05) high severity (3.29)	9%
Hopko et al 2005	N=6 (6)	None	Individual	9	Pre-, Post-treatment, 3-months	BDI-II, HDRS, CES-D	BDI-II (1.7, 2.46) HDRS	33%

							(1.7, 2.21)	
Porter et al, 2004	N=30 (30)	Wait-list control	Group	10	Pre-, Post- and 3-months	BDI-II, HDRS	BDI-II (.71, 1.17) HDRS (1.56, 1.72)	13%
Hopko et al 2003	N=25 (10)	Supportive psychotherapy	Individual	6	Pre-, post-treatment	BDI-II	BDI-II (1.56, na); differential effect with SP = .73	0% (inpatient)
Jacobsen et al 1996	N=152 (57)	CBT, AT	Individual	20	Pre-, post-, 6-, 12-, 18-, 24 mo	LIFE, BDI-II, HDRS	BDI (2.78, 2.92) HDRS (2.58, 2.48)	10%
Wright et al unpub'ed	N=62 (30)	Usual care	Group	8	Pre-, post-, 2 months	BDI, HDRS	BDI (.49, .30) HDRS (.93, .98)	20%

### *Recommendations for future research*

The present review reveals that BA for MDD meets criteria as an empirically supported intervention. More than two independent, well controlled investigations have been completed, all favoring BA over a controlled comparison, i.e. waitlist or alternative treatment or showing its equivalence to an existing frontline intervention. The attrition rates for BA are comparable on average to existing frontline interventions as well and better than rates achieved in most medication trials. The target populations so far have included rural low-SES sample, veterans sample with comorbid PTSD, a sample of MDD comorbid with cancer diagnoses, a sample of MDD comorbid with obesity diagnosis, and an undifferentiated sample of subjects recruited to a trial at an academic clinical research site. The quality of investigations has been moderate to very good in terms of gold standards for treatment outcome investigations. These facts notwithstanding, it is clear that additional large-scale trials are needed to establish confidence in this intervention as a front line treatment of choice. The following types of studies are recommended.

More studies are needed that compare BA with focused active psychological interventions. For example additional studies that compare BA to CBT or CT are necessary. This kind of investigation would essentially establish a cross-validation of BA with reference to the original Jacobson et al. study. Additional studies should use a common BA treatment manual chosen perhaps from among those demonstrating efficacy and stronger effect sizes with shorter intervention time periods (8-12 weeks), rather than the original Jacobson manual.

Studies should be conducted that compare BA to evidence-based medication interventions, alone and in combination with balanced lead-in periods (those starting with BA versus those starting with medications). Patient severity levels should be taken into account during random assignment to assure balance within samples.

Attention should be given to SES levels in patient comparisons, as a potential advantage of BA might be its relative simplicity and suitability to individuals without sophisticated verbal abilities as might be found among low SES individuals.

Given the promise of efficacy shown so far with BAGT, there should be additional studies with larger sample sizes that utilize this modality as a primary intervention, a follow along intervention, and a 'salvage' intervention during trials due to its potential cost effectiveness.

Continued study of BA in relation to comorbid medical and psychological conditions is also warranted, given the prevalence of MDD with such conditions as Posttraumatic Stress Disorder, anxiety disorders, and a vast array of primary medical conditions.

Finally, researchers and clinicians should not lose sight of the fact that BA's original emergence in the context of a dismantling study of CBT provides support for it as a parsimonious accounting as the active ingredient in CBT / CT. In this respect, BA might be considered the central ingredient to be heavily accentuated when attempting to increase the effective "dosage" of the full CBT/CT protocol. Research investigators and clinicians who address recalcitrant and unremitting MDD in the face of standard manualized CBT might consider giving greater emphasis to the BA component when working with these difficult patients.

### *Policy Implications*

BA is an empirically supported intervention for MDD. It should be practiced by trained and competent clinicians, and insurers should recognize it as an efficacious therapy for Major Depressive Disorder. As an emerging practice it warrants continued research by clinical investigators so that a determination can be made as to what types of patients respond best to this intervention in comparison to other efficacious treatments. NIMH's interest in this therapeutic intervention lays the groundwork for potential funding for larger scale studies.

### References

- Aikens, J., Nease, D., Nau, D., Klinkman, M., & Schwenk, T. (2005). Adherence to maintenance-phase antidepressant medication as a function of patient beliefs about medication. *Annals of Family Medicine* *Annals of Family Medicine*, 3(6) 1544-1709
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders – text revision* (4<sup>th</sup> ed.). Washington, DC.
- Center for the Advancement of Health. (2000). *Selected evidence for behavioral approaches to chronic disease management in clinical settings: depression*.
- Chambless, D. L., Baker, M. J., Baucom, D. H., Beutler, L. E., Calhoun, K. S., Crits-Christoph, P., et al. (1998). Update on empirically validated therapies, II. *The Clinical Psychologist*, 51(1), 3 – 16.
- Cullen, J.M., Spates, C. R., Pagoto, S. & Doran, N. (2006). Behavioral activation treatment for major depressive disorder: a pilot investigation. (Unpublished doctoral dissertation, Western Michigan University, 2006).
- Dimidjian, S. (2005). Behavioral activation, cognitive therapy, and antidepressant



- medication in the acute treatment of major depression. (Unpublished doctoral dissertation, University of Washington, 2005).
- Dobson, K. S. (1989). A meta-analysis of the efficacy of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 57*(3), 414-419.
- Foa, E. B. & Meadows, E. A. (1997). Psychosocial treatments for posttraumatic stress disorder: a critical review. *Annual Review of Psychology, 48*, 449-480.
- Gortner, E. T., Gollan, J. K., Dobson, K. S., & Jacobson, N. S. (1998). Cognitive-behavioral treatment for depression: relapse prevention. *Journal of Consulting and Clinical Psychology, 66*(2), 377-384.
- Hopko, D. R., Armento, M. E. A., Hunt, M. K., Bell, J. L., & Lejuez, C. W. (2005). Behavior therapy for depressed cancer patients in primary care. *Psychotherapy: Theory, Research, Practice, Training, 42*(2), 236-243.
- Hopko, D. R., Lejuez, C. W., & Hopko, S. D. (2004). Behavioral activation as an intervention for coexistent depressive and anxiety symptoms. *Clinical Case Studies, 3*(1), 37-48.
- Hopko, D. R., Lejuez, C. W., LePage, J. P., Hopko, S. D., & McNeil, D. W. (2003). A brief behavioral activation treatment for depression. *Behavior Modification, 27*(4), 458-469.
- Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., Gortner, E., & Prince, S. E. (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Counseling and Clinical Psychology, 64*(2), 295-304.
- Kimerling, R., Ouimette, P. C., Cronkite, R. C., & Moos, R. H. (1999). Depression and outpatient medical utilization: a naturalistic 10-year follow-up. *Annals of Behavioral Medicine, 21*, 317-321.
- Last, Cynthia (1985). Patterns of attrition for psychosocial and pharmacologic treatments of depression. *Journal of Clinical Psychiatry Vol 46*(9) (Sep 1985): 361-366
- Lewinsohn, P.M., Sullivan, J.M., & Grosscup, S.J. (1980). Changing reinforcing events: An approach to the treatment of depression. *Psychotherapy: Theory, Research, and Practice, 47*, 322-334.
- Mulick, P. S. & Naugle, A. E. (2004). Behavioral activation for comorbid PTSD and major depression: a case study. *Cognitive and Behavioral Practice, 11*(4), 378-387.
- National Institute of Mental Health (2005). Breaking Ground, Breaking Through: Strategic Plan for Mood Disorders Research. U.S. Department of Health & Human Services.
- Porter, J. F., Spates, C. R., & Smitham, S. (2004). Behavioral activation group therapy in public mental health settings: a pilot investigation. *Professional Psychology: Research and Practice, 35*(3), 297-301.
- Satterfield, Jason M. (1998). Cognitive behavioral group therapy for depressed, low-income minority clients: Retention and treatment enhancement. *Cognitive and Behavioral Practice Vol 5*(1) 65-80

Wright, T. P., Spates, C. R., Bloem, W. D., DeViva, J. C., & Pope, S. A. (2003). The effectiveness of behavioral activation group therapy: treating comorbid depression on a specialized inpatient posttraumatic stress disorder unit for combat veterans. Presented at the 37<sup>th</sup> annual conference of the Association for the Advancement of Behavior Therapy, Boston, MA.

World Health Organization. (2006). WHO Depression. Retrieved July 20<sup>th</sup>, 2006 from [http://www.who.int/mental\\_health/management/depression/definition/en/](http://www.who.int/mental_health/management/depression/definition/en/)

**Author Contact Information:**

**C. Richard Spates, Ph.D. (Corresponding author)**  
Western Michigan University  
Department of Psychology  
1903 West Michigan Avenue  
Kalamazoo, MI 49008  
Tel: 269 387-4329  
e-mail: [rspates@wmich.edu](mailto:rspates@wmich.edu)

**Sherry L. Pagoto, Ph.D.**  
Assistant Professor of Medicine  
Department of Medicine  
Division of Preventive and Behavioral Medicine  
University of Massachusetts Medical School  
55 Lake Avenue North  
Worcester, MA 01655  
e-mail: [sherry.pagato@umassmed.edu](mailto:sherry.pagato@umassmed.edu)

**Alyssa Kalata, B.A.**  
Western Michigan University  
Department of Psychology  
1903 West Michigan Avenue  
Kalamazoo, MI 49008  
e-mail: [alyssa.h.kalata@wmich.edu](mailto:alyssa.h.kalata@wmich.edu)