Published in final edited form as:

Behav Modif. 2010 July ; 34(4): 310–329. doi:10.1177/0145445510373384.

Measuring Homework Completion in Behavioral Activation

Andrew M. Busch^{1,2}, Lisa A. Uebelacker^{1,3}, Zornitsa Kalibatseva⁴, and Ivan W. Miller^{1,3}

¹Warren Alpert Medical School of Brown University

²The Miriam Hospital

³Butler Hospital

⁴Michigan State University

Abstract

The aim of this study was to develop and validate an observer-based coding system for the characterization and completion of homework assignments during Behavioral Activation (BA). Existing measures of homework completion are generally unsophisticated, and there is no current measure of homework completion designed to capture the particularities of BA. The tested scale sought to capture the type of assignment, realm of functioning targeted, extent of completion, and assignment difficulty. Homework assignments were drawn from 12 (mean age = 48, 83% female) clients in two trials of a 10-session BA manual targeting treatment-resistant depression in primary care. The two coders demonstrated acceptable or better reliability on most codes, and unreliable codes were dropped from the proposed scale. In addition, correlations between homework completion and outcome were strong, providing some support for construct validity. Ultimately, this line of research aims to develop a user-friendly, reliable measure of BA homework completion that can be completed by a therapist during session.

Keywords

Behavioral Activation; homework completion

Behavioral Activation (BA; Lejuez, Hopko, & Hopko, 2001; Martell, Addis, & Jacobson, 2001) is a well-established treatment for depression (Coffman, Martell, Dimidjian, Gallop, & Hollon, 2007; Dimidjian et al., 2006) and is now being applied to a wide variety of populations (e.g., Daughters et al., 2008; Gawrysiak, Nicholas, & Hopko, 2009). The central focus of BA is very simple: to increase behaviors that put patients in contact with diverse and stable sources of positive reinforcement in their environment (Kanter, Busch, & Rusch, 2009). BA seeks to do this by scheduling out-of-session assignments that are active, pleasurable, and/or goal-directed (including unpleasant responsibilities that the patient has been avoiding), which are broadly termed *activation assignments*. If such activation occurs, BA theory posits that patients will experience increased positive reinforcement and depressed mood will decrease. Put sequentially, BA's mechanism of change is as follows: completion of activation assignments → increase in the patient's overall rate of pleasurable

© The Author(s) 2010

Corresponding Author: Andrew Busch, Post-doctoral Fellow, Centers for Behavioral and Preventive Medicine, Brown Medical School, Coro Bldg West, Suite 500, One Hoppin St, Providence, RI 02903, Andrew_busch@brown.edu.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

and goal directed behavior → increased contact with positive reinforcement → mood improvement.

As interest in BA has risen over the past decade, researchers have sought to develop and validate measures that capture BA's mechanism of change. Most notably, Kanter and colleagues have developed the BA for Depression Scale (BADS; Kanter, Mulick, Busch, Berlin, & Martell, 2007; Kanter, Rusch, Busch, & Sedivy, 2009) and Armento and Hopko (2007) have developed the Environmental Reward Observation Scale (EROS). Both scales were designed to capture the process of BA across the course of treatment and are complementary in that they target different points in the causal chain that makes up BA's mechanism. The BADS measures the patient's overall level of activation and the EROS measures the patient's current level of contact with positive reinforcement. Thus, established and validated measures exist that capture two components of BA's mechanism (i.e., overall rate of pleasurable and goal-directed behavior and contact with positive reinforcement). However, measurement of activation assignment completion in BA has thus far been limited to yet-to-be validated patient report diary cards (e.g., "behavior check-out" included in the Lejuez et al.'s, 2001, treatment manual) and single-item therapist ratings of completion on the session level (e.g., Addis & Jacobson, 2000).

The lack of development of sophisticated measures that assess activation assignment completion is surprising, given the centrality of homework in BA. In fact, one could argue that all BA techniques are explicitly aimed at identifying relevant between-session homework targets (i.e., self-monitoring, goal/value assessments), assigning homework (i.e., activity scheduling), or increasing the likelihood of successful homework completion (e.g., social skills training). Capturing rate of activation homework completion would provide the most immediate measure of the effects of in-session BA techniques and would fill a gap left by existing BA process measures.

There is a wealth of research suggesting that homework completion is related to positive outcomes in psychotherapy in general (Kazantzis, Deane, & Ronan, 2004) and in cognitive-behavioral therapy for depression in particular (e.g., Burns & Nolen-Hoeksema, 1992). Thus, the possibility of applying homework-related measures used in previous studies to BA assignments was carefully considered. Unfortunately, measures employed in past studies have lacked sophistication. Specifically, many utilized potentially unreliable methods of measurement, such as relying on patient or therapist retrospective report over multiple sessions, failed to capture the proportion of homework completed, and/or presented little or no data on rating reliability (Detweiler & Whisman, 1999; Primakoff, Epstein, & Covi, 1986). Many major studies in this area have relied on single-item measures of homework completion (e.g., Burns & Spangler, 2000), which precludes any sophisticated analysis of completion itself.

An additional problem with applying existing measures of homework completion to BA assignments is that these measures are insensitive to the idiosyncrasies of BA's purported mechanism of change. In BA, patient out-of-session increases in activation are paramount, but not all BA homework assignments relate to activation equally. For example, self-monitoring assignments ask patients to simply observe rather than change their out-of-session behavior. Self-monitoring assignments may be an important component of BA treatment, but it may also be useful to analyze these assignments separately. Furthermore, recent BA manuals have become more focused on the importance of activation assignments being based on the patient's values (Gaynor & Harris, 2008; Uebelacker, Weisberg, Haggarty, & Miller, 2009). Implicitly, this movement toward value-based BA indicates that assignments in areas of greatest value to the patient should have a greater effect on outcome than assignments in less meaningful areas. Though the importance of value-related

activation has yet to be questioned empirically, in order for a BA homework completion rating system to be sensitive to this issue, information on the life area or realm of functioning targeted by the activation assignment is necessary.

The current study sought to develop and validate an observer-based coding system for the characterization and completion of BA homework on the level of individual homework tasks. Specifically, we sought to capture the type of assignment, the realm of functioning targeted, assignment difficulty, and extent of completion. The ultimate goal of this line of research is to develop a user-friendly, reliable measure of homework completion that can be used by BA therapists to rate completion of patient homework during or following each session.

Method

Homework assignments for analysis were drawn from patients in a completed open trial (Trial 1; Uebelacker et al., 2009) and an ongoing randomized trial (Trial 2, K23 MH067779) of a 10-session manual of BA for treatment-resistant depression in a primary care setting. Treatment in both studies was conducted by the same two therapists from 2006 to early 2009. One therapist was an advanced clinical psychology graduate student during Trial 1 and a PhD during Trial 2. The second therapist was a licensed PhD-level clinician at the time of both trials.

The BA manual utilized in these trials was specifically tailored to a primary care setting and to encourage collaboration with primary care physicians (Uebelacker et al., 2009). These adaptations included conducting treatment in a primary care clinic, emphasizing physical health behaviors in activation assignments, monitoring and encouraging adherence to anti-depressant medication regimens, and limiting the number of sessions. Providing assignments that were consistent with the patient's values and long-term goals was also emphasized. In these trials, each patient could attend up to 10 sessions over 4 months. Further details of this treatment manual can be found in Uebelacker and colleagues (2009).

Data from these trials provide a unique opportunity for development of an observer-based behavioral homework completion rating scale because audiotapes of sessions were available for multiple patients in Trial 1 and the majority of patients in Trial 2 and the treatment manual structured homework assignment and review. Specifically, the manual directed therapists to (a) assign activation homework during every session following agreement on treatment goals (usually this agreement occurred by the third session), (b) record all assignments given in clinical notes, and (c) review homework compliance at the beginning of every session. Thus, raters had easy access to all homework-related information to which the therapists had access.

Patients who at the time of the current study had attended at least 6 of 10 possible sessions were considered for inclusion. Inclusion also required that between Session 4 and Session 7, at least two sessions met the following criteria: (a) audio recordings were available to coders, (b) recordings were of sufficient audio quality for coding, (c) at least one task was assigned during the previous session, (d) there were no more than 21 days since the previous session, and (e) the patient attended session alone. If both Sessions 4 and 6 met the criteria discussed earlier, then they were used for analysis. If either 4 or 6 did not meet all of the criteria discussed, the closest available session (either 5 or 7) was used to replace it.

Only Sessions from 4 to 7 were considered for analysis because many patients were not assigned homework in the first few sessions (as treatment goals were not yet set), and many patients did not attend all 10 sessions available to them. A maximum of 21 days between sessions was imposed to limit patient forget-fulness of homework completion. Sessions

where family members attended were excluded because such attendance may have prompted agenda shifts that left less time for homework review.

Of the 12 patients in Trial 1, 10 attended at least 6 sessions. Of these 10, the sessions of 4 patients had been regularly audio-taped. One of these 4 was excluded due to poor audio quality of relevant sessions and 1 was excluded due to the length of time between sessions. Thus, 2 patients in the current study were pulled from Trial 1.

At the time of the current study, 24 patients had begun treatment in Trial 2. Seventeen of these patients had completed at least 6 sessions. An additional 7 patients were excluded from the current study due to missing tapes from relevant sessions ($n = 4$), poor audio quality ($n = 1$), no assignments given prior to relevant sessions ($n = 1$), and the length of time between relevant sessions ($n = 1$). Thus, 10 patients in the current study were pulled from Trial 2. Of these 10 patients, 4 patients received BA only after crossing over from a 4-month waitlist control condition.

Inclusion requirements stated earlier yielded 12 eligible patients (M age = 47.5, 83.3% female), producing 24 sessions for analysis (10 from Session 4, 4 from Session 5, 8 from Session 6, and 2 from Session 7). A total of 82 discrete homework assignments were given in the sessions before those chosen for analysis (Mean assignments per session = 3.42, range 1–7) as reported by the therapist in the clinical note for the previous session. Before any coder training, Coder 2 (a postbachelor's research assistant) listened to audiotapes of all 24 sessions in their entirety and transcribed all session segments that were in any way relevant to the completion of homework.

Development of a Coding System

Development of this rating system began with the recommendations of Primakoff and colleagues (1986) who provided direction on measuring homework adherence. Their main recommendations were that homework rating scales should (a) produce ratings on the level of individual sessions (rather than estimates across sessions), (b) address the portion of the assignment completed, and (c) address the time spent on the assignment. The system tested in this study is on the individual assignment level and specifically sought to determine percent of assigned tasks completed and time spent on homework tasks. Primakoff and colleagues also argued for the inclusion of quality of homework ratings. However, their rationale for this recommendation was based on the assumption that the content of assignments is similar across patients, which is not true in BA. Thus, the quality of assignment completion was not rated in this study.

To develop the coding system, initial items, response options, and format were rationally derived via discussions between the first and second author. Next, clinical notes of patients in the two trials mentioned earlier who did not meet inclusion criteria were reviewed by the first author (also Coder 1). This review provided a sense of the range of both assignment content and patient completion of assignments, which were incorporated into items. At this point, a BA expert, otherwise not involved in this study, reviewed this draft system and suggested changes. Finally, Coder 1 and Coder 2 attempted to apply this evolving system collaboratively to transcripts of several session segments (again pulled from patients in the two trials who did not meet inclusion criteria) and to several hypothetical assignment completion scenarios. This process yielded the coding system detailed in the following section.

Coding System

Codes included both those meant to characterize the assignments given (i.e., "assignment characteristic" codes) and to detail the extent to which those assignments were completed

(i.e., “assignment completion” codes). The following assignment characteristics were coded: Assignment type (see Appendix, Item 1), realm of functioning (see Appendix, Item 2), quantifiability (see Appendix, Item 3), and assignment difficulty.

The assignment type code was explicitly designed to categorize assignments in such a way that the isolation of assignments most closely related to behavior change (Options A and B) from other types of assignments would be possible. Addition of the PRN response option (Option C) was in recognition of recent findings, suggesting that one of the most common reasons for uncompleted homework tasks was that no occasion for homework completion occurred between sessions (Helbig & Fehm, 2005). This allows for differential treatment of assignments that are conditional on the occurrence of other between-session events that may or may not occur. The self-monitoring response option (Option D) was included because self-monitoring of relevant behavior and mood is a core assignment in all forms of BA but does not involve explicit behavior change. The abstinence/limitation response option (Option E) was added during the development of the system after it was noted by coders that some assignments were negatively worded (e.g., limit time watching TV to 1 hour per day).

The realm of functioning code was designed to allow for future explorations of two issues. First, this code would allow for tests of possible differential relationships between areas of activation and mood. For example, one may hypothesize that completion of social activities (Options A and B) would lead to immediate improvements in mood, whereas catching up on professional or academic tasks and household chores (Options C and F) would lead to later mood improvements. The second issue is similar to the first, but on the patient level. Specifically, this code allows for an exploration of the match between the patient’s idiographic values and life goals and the homework he or she is completing. Consider a patient who expresses that her paramount value is to improve her relationship with her spouse and family. One may hypothesize that she would experience more positive outcomes if she were highly vigilant of completion of family-related assignments rather than neglectful of family assignments when completing the majority of assignments in other areas.

The quantifiability code was added during measure development in response to the observation by coders that most assignments were not framed in terms of easily quantifiable details (i.e., time, distance, amount, or frequency). Furthermore, it was observed that some assignments, while specific and discrete, would be difficult to split into partial completion (e.g., call your doctor to reschedule an appointment).

Assignment difficulty was rated on a 7-point scale, including anchors such as *not at all*, *moderately*, and *extremely*, with example assignments for each anchor. Coders were instructed to code difficulty on the basis of how hard the assignment would be for an average nondepressed person. This was done because the coders were not familiar with the history or course of treatment of individual patients. Difficulty of assignments was coded to explore the possibility of giving completion of difficult assignments more weight than relatively easy assignments.

Assignment completion codes were categorical completion (i.e., completion on an ordinal scale; see Appendix, Item 4), percentage completion (i.e., completion on a ratio scale; see Appendix, Item 5), and time spent on assignment.

When coding either categorical or percentage completion of an activity or repeated activity, coders were allowed to consider substituted activities that were not exactly what was assigned. Coders were instructed to allow a substitute activity only if it was highly functionally similar to the assigned activity. This option was added in response to the criticism raised by Kazantzis and Lampropoulos (2002) that past scales have not accounted

for the possibility that patients may engage in between-session activities that are only slightly different from those assigned. The exemplar provided for this substitution was a patient who was assigned to walk on the treadmill at the gym for half an hour, but instead chose to go for a half-hour bike ride because the weather was nice. Option C for the categorical completion code (“Made an attempt or effort to start assignment”) was included because it was recommended by Primakoff and colleagues (1986). An illustrative example would be a patient who was assigned to borrow a book from the library between sessions, who actually made it to the library, but left before borrowing a book because it was too busy. On the percentage completion code, if a patient completed more than 100% of an assignment (e.g., patient exercised 4 times when only assigned to do so 3 times) coders could assign values greater than 100%.

Time spent was recorded in minutes whenever any mention of time spent on the assignment was made by the patient. When the patient reported not making any progress on an assignment, time spent was recorded as zero.

Coding Procedure

Before coding began, coders collaboratively identified assignments that were repeats of previous assignments for the same patient (e.g., one patient was asked to “go for an early morning walk twice per week” after both the third and fifth sessions). These repeated assignments were only coded once for assignment characteristic purposes, in order to avoid double-counting agreement or disagreements when calculating reliability. Coding of assignments was completed in two passes by two coders. Coder 1 was an advanced PhD graduate student, and Coder 2 was a postbachelor’s research assistant at the time of coding. First, assignment characteristics were coded on the basis of information provided in the clinical note for the previous session. Then, transcripts of homework completion discussions were reviewed to code homework completion variables.

Each coder first completed both assignment characteristics and assignment completion coding independently. Following independent coding, disagreements on all codes other than assignment difficulty and percent completion were discussed between coders until a consensus code was agreed upon. The average of the two ratings was considered the consensus code for percentage completion and assignment difficulty. All results reported in the following section, other than those regarding reliability, are based on these consensus codes.

Analysis

Reliability was examined between coders for all variables. Cohen’s Kappas were calculated for all nominal variables, and intraclass correlations were calculated for all continuous variables. Frequencies of nominal and ordinal codes will be presented and distributions of continuous items will be described.

Validation of a homework completion scale and whether that scale predicts outcome are separate issues. However, because it is well established that overall homework completion is associated with greater symptom reduction, a positive association between a valid homework completion measure and treatment outcome would be expected. Meta-analyses have estimated the correlation between homework completion and outcome to be .22 (Kazantzis, Deane, & Ronan, 2000). Observing a similar or larger effect size in this study can serve as an indicator of construct validity. Thus, the relationship between homework completion and change in depression symptoms was examined. Homework completion was aggregated to the patient level using multiple metrics. First, means for both percentage completion and categorical completion codes across both sessions were calculated for each

patient. This method indicates how much of the homework assigned a patient completed. Second, total number of assignments a patient completed was calculated by summing all percentage completion responses for each patient and by a simple frequency count of total assignments coded “fully” and “partially” completed on the categorical completion code for each patient. This second method of aggregation speaks to the number of homework assignments the patient completed across the two sessions, regardless of how many were assigned.

Treatment response was defined by change in Quick Inventory of Depression Symptoms-Clinician Rating (QIDS-C; Rush et al., 2003) scores from pre- to post-treatment assessment. The QIDS is a 16-item structured interview assessing depression severity that is based on *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV*; 4th ed., American Psychiatric Association, 1994). The QIDS has been found to have good psychometric properties (Rush et al., 2003) and has been utilized in large scale depression treatment trials.

Results

Reliability

Assignment characteristics—Of the 82 homework assignments analyzed, 14 were assigned to the same patient twice. Thus, 68 unique assignments were subjected to assignment characteristic coding. Table 1 presents agreement percentage and intercoder kappa for assignment type, realm of functioning, and quantifiability and the intraclass correlation for assignment difficulty. All kappas are at or above an acceptable level of agreement (Fleiss, 1981). The intraclass correlation for ratings of assignment difficulty was below that of expected for reliable coding.

Assignment completion—Of the 82 homework assignments considered, 9 were not reviewed by the therapist and no mention was made of assignment completion. Thus, 73 assignments were subjected to completion coding. Table 2 presents agreement percentage and Cohen’s kappa for categorical completion and intraclass correlations for percentage completion and time spent. Data are presented for all assignments reviewed by the therapist, for those assignments coded as either “quantifiable” or “partially quantifiable”, and the nine assignments rated as “unquantifiable.” This was done because, as expected, a large decrease in reliability was observed for those nine assignments coded as “unquantifiable.” All Kappas were in the acceptable range. Note that the *N*’s for time spent are low because duration of assignment follow-through was not mentioned for the majority of assignments.

Frequency and Distribution of Codes

Table 3 presents frequency and percentage of total codes for assignment type, realm of functioning, quantifiability, and categorical completion. Note that all data presented in Table 3 reflect the consensus codes.

The distribution of the percentage completion code on the level of individual homework assignments was highly bimodal, 22 out of 73 (30.14%) assignments were coded as 0% completed, and 26 out of 73 (35.62%) assignments were coded as 100% completed. However, when averaged to the session level, the distribution was relatively normal ($M = 58.23\%$, $SD = 26.83\%$, $Skewness = -.15$, $Kurtosis = -.39$). Regarding the time spent code, only six assignments were given a nonzero code by either coder. This resulted in a greatly negatively skewed distribution.

Relation to Outcome

Each patient's average percentage completion and categorical completion scores were calculated and related to outcome (change in QIDS scores across treatment), producing Pearson correlations of .47 (nonsignificant) and .55 ($p = .06$), respectively. In an alternative scoring system, percentage completion scores for each item within patient were summed and categorical completion responses coded as "partially" or "fully" completed where counted within patient. This method produced correlations with outcome of .42 (nonsignificant) for percentage completion and .51 ($p = .09$) for categorical completion. All analyses relating homework completion to outcome were also completed after removing the nine assignments that were coded as unquantifiable, but results were highly similar to those reported earlier.

Discussion

These results demonstrate that most aspects of homework assignment and completion relevant to BA can be reliably coded, using only information available to the therapist, by coders who are unfamiliar with the patient. As for assignment characteristics, coders were able to reliably identify assignment type, realm of functioning, and quantifiability but were unable to code assignment difficulty reliably. Regarding homework completion, coders were able to reliably code categorical completion and percentage completion. Though technically coders demonstrated adequate reliability when coding time spent, problems with this variable (discussed later) limit its usefulness in an observer-rated coding system. The observed relationship between homework completion items and outcome was consistent with earlier findings and provides some construct validity to this measure.

Although no inferential analyses were conducted on assignment type and realm of functioning in the current study, these variables both have potential to be important in future studies of BA homework completion. Regarding assignment type, 9% of assignments were identified as differing from the stereotypical BA assignment of engaging in single or multiple activities. It should be noted that self-monitoring assignments are given more often in early BA sessions; thus, the percentage of assignments coded as self-monitoring would probably be greater in studies including sessions from the entire course of treatment.

It may be useful to analyze assignments in these three nonactivity categories separately or to develop more nuanced algorithms for calculating completion. For example, in the current study one patient was assigned to "call her grandchildren on the phone if they did not come by during the week." However, her grandchildren visited her between sessions. Thus, though she did not complete the assignment, completion was not necessary. This assignment was considered a PRN assignment and was coded as 0% completed. Creation of subcodes that indicate whether an assignment was not completed due to lack of opportunity would allow for a more valid treatment of this situation.

Self-monitoring and abstinence/limitation assignments create similar problems in terms of measuring BA's mechanism of change. Would one expect successful monitoring of time watching television to affect mood in the same manner or to the same extent that successfully attending a social event would? Furthermore, the relationship between activation and abstinence/limitation assignments is highly dependent on what replacement behavior the patient engaged in. Consider a patient who was assigned to reduce television viewing from 4 hours per day to 1 hour per day. If the patient completes the assignment but fills those 3 hours with another inactive behavior (e.g., rumination, playing video games alone), would one expect this patient to become less depressed? Although these are empirical questions, a highly valid measure of activation homework completion would take these issues into account and this will be a focus in follow-up studies.

Regarding realm of functioning, the success of this code will allow for future explorations of how a targeted life area interacts with mood changes on the group and individual level. Exploring the relationship between value consistency and depression change may be particularly fruitful. Although not considered at the time of the design of the current study, two recent publications speak to the potential usefulness of this code. First, Hopko and Mullane (2008) found that individuals experiencing mild depression engaged in fewer behaviors in certain life areas; specifically, they spent less time engaging in social, physical, and educational activities. However, they found no differences in engagement in several behaviors that one might expect such differences (i.e., health and hygiene), suggesting that tracking the life areas where the patient's behavior is changed through homework completion may be useful. Second, realm of functioning items in the current study are highly similar to those in the recently published Valued Living Questionnaire (Wilson, Sandoz, Kitchens, & Roberts, in press). In future studies, it may be useful to modify response options in the realm of functioning code slightly to be more consistent with this measure and to facilitate their parallel and complementary use.

The quantifiability of assignments as recorded by therapists appears to affect coding reliability. This variable is really a measure of how explicit therapists were regarding what is expected of patients between sessions. An illustrative example of an assignment coded as "unquantifiable" was "be mindful of diet as it relates to your diabetes management." There is nothing inherently wrong with this assignment, and it was clearly linked to a valued life goal. However, the coders had very limited perspective to determine whether the assignment has been completed and disagreement occurred. Though giving general or unspecified assignments such as this may be clinically useful at times, the current study suggests that they should be excluded from analysis when observer-based measures of homework completion are utilized.

It should be noted that this study was not planned until all sessions included in these analyses had been completed, and the study therapists were unaware that any post hoc homework coding would occur. Thus, it is possible that the problem of "unquantifiable" assignments may be mitigated by instructing therapists to word assignments in more measurable ways. For example, if the earlier stated assignment were rephrased as "be mindful of your diabetes management by keeping a daily food diary," it could easily be reliably coded. Given the long-held recommendation that therapy homework assignments should be as specific and concrete as possible (Shelton & Levy, 1981), the quantifiability code could be used as a simple measure of therapist adherence to this recommendation. It is unknown how distribution of this code in the current sample would relate to other trials of BA or trials of other treatments for depression. Overall, this study suggests that quantifiability might be an important variable to consider when assigning, recording, and measuring completion of behavioral homework assignments.

Inability to reliably code assignment difficulty was particularly disappointing, as we had hoped to explore methods for capturing the interaction between quantity and difficulty of homework completed. It should be noted that though observers unfamiliar with a patient's full course of treatment, as in the case for the coders in this study, may not be able to reliably code homework difficulty, it is possible that the therapist or patients themselves (Kazantzis et al., 2004) would be able to make difficulty ratings in a reliable and valid manner. Thus, although we have not included any difficulty rating in our proposed observer-rated measure in the Appendix, a possible parallel measure completed by therapists during or after sessions could incorporate some sort of valid difficulty rating.

Levels of coder reliability for categorical completion and percentage completion are quite promising. When unquantifiable assignments were excluded, reliability statistics for both

variables were acceptable. However, percentage completion has the advantage of demonstrating high reliability even when unquantifiable assignments were included. These results suggest that valid prospective and post hoc evaluations of homework completion could be conducted on large-scale trials of BA with additional effort. However, it should be noted that these items suffer from the same potential biases of all retrospective self-report items and have yet to be compared to ratings made by independent observers (e.g., partner report) or data collected through more objective methods (e.g., ecological momentary assessment).

Although the time spent variable was reliably coded, inspection of the distribution of this code suggests that it was invalid. Specifically, only six assignments were coded as nonzero by either coder. That is, time spent on assignment follow-through was only mentioned during 6 of the 73 assignments reviewed. All other coder responses to this item were automatically zeros because the assignment had been rated as 0% completed. Given that actual time information was available for less than 10% of assignments, we have chosen to leave the time spent variable off of our proposed measure in the Appendix. Again, in a study that was planned a priori to assess homework variables, this issue may be easily solvable by asking therapists specifically to ask patients about how much time was spent on each assignment. However, in post hoc studies, such as the current one, coding for time spent on homework does not appear to be viable.

The degree of association observed between metrics of homework completion and treatment outcome in this study is stronger than that reported in past meta-analytic studies (.22; Kazantzis et al., 2000). However, with the small sample used in the current study, none of these relationships reached statistical significance at the .05 level. Nonetheless, these correlations provide this scale with some construct validity. The two methods of aggregating completion items (proportion of assigned homework completed and total number of assignments completed) produced highly similar associations with outcome. Thus, either method could be utilized in future investigations. Note that conclusions regarding BA's mechanism of change that can be drawn from these results are limited for multiple reasons. First, no temporal precedence was established in this study. That is, measurement of homework completion did not precede measurement of depression change. Second, patients in this study were also receiving concurrent antidepressant medication. One would expect that some variance in depression change would be related to medication effects, thus complicating any conclusions regarding mechanism.

This study is somewhat limited by its small sample (24 sessions drawn from 12 patients). However, it is a successful first step in a line of research with the end goal of developing and validating a reliable and user-friendly measure of homework completion in BA. The current measure could be applied to ongoing or completed trials of BA to assess for the first step in BA's mechanism of change occurring between individual sessions. In future studies, we hope to apply a lagged correlation method to demonstrate that completion of homework is reflected in increases of general activation, subsequent increases in environmental reinforcement, and eventual decreases in depressed mood.

Although the current measure was designed and tested for use by independent observers, a parallel therapist version would have several advantages. Most important, the therapist has the opportunity to ask the patient follow-up questions, whereas a coder does not. If the therapist completed the measure during or immediately following session, he or she could specifically query during homework review regarding relevant items (e.g., "What percentage of your house cleaning would you say you completed this week?"). Furthermore, a therapist may be able to reliably rate homework difficulty or similar variables ideographically for specific patients (e.g., "How difficult is this assignment for this client to complete given

their current functioning?”). Finally, if therapists were to complete this measure in session, it may help guide therapist assignment and review of homework to be more BA consistent. Future studies will focus on the development of such a parallel therapist measure.

Acknowledgments

The authors would like to thank David Baruch, Jonathan Kanter, and Brandon Gaudiano for their useful comments on earlier versions of this manuscript.

Financial Disclosure/Funding

This study was supported by a grant from the NIMH to Dr. Uebelacker (K23 MH067779) and an intramural grant from the Brown Medical School Clinical Psychology Training Consortium to Dr. Busch.

Biographies

Andrew M. Busch is a postdoctoral fellow at the Centers for Behavioral and Preventative Medicine at the Alpert Medical School of Brown University. His research interests include behavioral treatments for depression and the adaptation of behavioral activation for novel populations.

Lisa A. Uebelacker is an assistant professor (research) of psychiatry and human behavior and an assistant professor of family medicine at Alpert Medical School of Brown University. She is also a research psychologist at Butler Hospital in Providence, RI. Her research interests focus on behavioral treatments for depression.

Zornitsa Kalibatseva is a graduate student in the Department of Psychology at Michigan State University. She is interested in studying cross-cultural aspects of psychopathology and especially how culture may affect the experience, expression, and treatment of depression.

Ivan W. Miller is a professor of psychiatry and human behavior of psychiatry and human behavior at Alpert Medical School of Brown University and director of the psychosocial research program at Butler Hospital in Providence, RI. His research interests include assessment and treatment of mood disorders, with specific interests in the treatment of suicidal patients and in family approaches to mood disorders.

References

- Addis ME, Jacobson NS. A closer look at the treatment rationale and homework compliance in cognitive-behavioral therapy for depression. *Cognitive Therapy and Research* 2000;24:313–326.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed.. Washington, DC: Author; 1994.
- Armento MEA, Hopko DR. The Environmental Reward Observation Scale (EROS): Development, validity, and reliability. *Behavior Therapy* 2007;38:107–119. [PubMed: 17499078]
- Burns DD, Nolen-Hoeksema S. Therapeutic empathy and recovery from depression in cognitive-behavioral therapy: A structural equation model. *Journal of Consulting and Clinical Psychology* 1992;60:441–449. [PubMed: 1619098]
- Burns DD, Spangler DL. Does psychotherapy homework lead to improvements in depression in cognitive-behavioral therapy or does improvement lead to increased homework compliance? *Journal of Consulting and Clinical Psychology* 2000;68:46–56. [PubMed: 10710839]
- Coffman SJ, Martell CR, Dimidjian S, Gallop R, Hollon SD. Extreme nonresponse in cognitive therapy: Can behavioral activation succeed where cognitive therapy fails? *Journal of Consulting and Clinical Psychology* 2007;75:531–541. [PubMed: 17663608]
- Daughters SB, Braun AR, Sargeant MN, Reynolds EK, Hopko DR, Blanco C, Lejuez CW. Effectiveness of a brief behavioral treatment for inner-city illicit drug users with elevated depressive

- symptoms: The life enhancement treatment for substance use (LETS Act!). *Journal of Clinical Psychiatry* 2008;69:122–129. [PubMed: 18312046]
- Detweiler JB, Whisman MA. The role of homework assignments in cognitive therapy for depression: Potential methods for enhancing adherence. *Clinical Psychology: Science and Practice* 1999;6:267–282.
- Dimidjian S, Hollon SD, Dobson KS, Schmalting KB, Kohlenberg RJ, Addis ME, et al. 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* 2006;74:658–670. [PubMed: 16881773]
- Fleiss JL. Balanced incomplete block designs for inter-rater reliability studies. *Applied Psychological Measurement* 1981;5:105–112.
- Gawrysiak M, Nicholas C, Hopko DR. Behavioral activation for moderately depressed university students: Randomized controlled trial. *Journal of Counseling Psychology* 2009;56:468–475.
- Gaynor ST, Harris A. Single-participant assessment of treatment mediators: Strategy description and examples from a behavioral activation intervention for depressed adolescents. *Behavior Modification* 2008;32:372–402. [PubMed: 18391053]
- Helbig S, Fehm L. Problems with homework in CBT: Rare exception or rather frequent? *Behavioural and Cognitive Psychotherapy* 2005;32:291–301.
- Hopko DR, Mullane CM. Exploring the relation of depression and overt behavior with daily diaries. *Behaviour Research and Therapy* 2008;46:1085–1089. [PubMed: 18687422]
- Kanter, JW.; Busch, AM.; Rusch, LC. *Behavioral activation: Distinctive features*. London: Routledge; 2009.
- Kanter JW, Mulick PS, Busch AM, Berlin KS, Martell CR. The Behavioral Activation for Depression Scale (BADs): Psychometric properties and factor structure. *Journal of Psychopathology and Behavioral Assessment* 2007;29:191–202.
- Kanter JW, Rusch LC, Busch AM, Sedivy SK. Validation of the Behavioral Activation for Depression Scale (BADs) in a community sample with elevated depressive symptoms. *Journal of Psychopathology and Behavioral Assessment* 2009;31:36–42.
- Kazantzis N, Deane FP, Ronan KR. Homework assignments in cognitive and behavioral therapy: A meta-analysis. *Clinical Psychology: Science and Practice* 2000;7:189–202.
- Kazantzis N, Deane FP, Ronan KR. Assessing compliance with homework assignments: Review and recommendations for clinical practice. *Journal of Clinical Psychology* 2004;60:627–641. [PubMed: 15141396]
- Kazantzis N, Lampropoulos GK. Reflecting on homework in psychotherapy: What can we conclude from research and experience? *Journal of Clinical Psychology* 2002;58:577–585. [PubMed: 11967882]
- Lejuez CW, Hopko DR, Hopko SD. A brief behavioral activation treatment for depression: Treatment manual. *Behavior Modification* 2001;25:255–286. [PubMed: 11317637]
- Martell, CR.; Addis, ME.; Jacobson, NS. *Depression in context: Strategies for guided action*. New York, NY: Norton; 2001.
- Primakoff L, Epstein N, Covi L. Homework compliance: An uncontrollable variable in cognitive therapy outcome research. *Behavior Therapy* 1986;17:433–446.
- Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-item Quick Inventory of Depressive Symptomatology (QIDS), clinician ratings (QIDS-C), and self-report (QIDS-SR): A psychometric evaluation patients with chronic major depression. *Biological Psychiatry* 2003;54:573–583. [PubMed: 12946886]
- Shelton JL, Levy RL. A survey of the reported use of assigned homework activities in contemporary behavior therapy literature. *Behavior Therapist* 1981;4:13–14.
- Uebelacker LA, Weisberg RB, Haggarty R, Miller IW. Adapted behavior therapy for persistently depressed primary care patients: An open trial. *Behavior Modification* 2009;33:374–395. [PubMed: 19282506]
- Wilson KG, Sandoz EK, Kitchens J, Roberts ME. The Valued Living Questionnaire: Defining and measuring valued action within a behavioral framework. *Psychological Record*. (in press).

Table 1

Percent Agreement and Inter coder Kappas of Assignment Characteristic Code

	% Agreement	χ	Intraclass r
Assignment type	91.18%	.84	
Realm of functioning	88.24%	.86	
Quantifiability	88.24%	.74	
Assignment difficulty			.37

Table 2

Assignment Completion Codes Stratified by Quantifiability

	% Agreement	χ	Intraclass r
Categorical completion			
All reviewed assignments ($n=73$)	76.71	.65	
Only assignments rated as fully or partially quantifiable ($n = 64$)	83.56	.72	
Unquantifiable assignments ($n = 9$)	44.44	.10	
Percentage completion			
All reviewed assignments ($n = 73$)			.79
Only assignments rated as fully or partially quantifiable ($n = 64$)			.82
Unquantifiable assignments ($n = 9$)			.52
Time spent			
All available ($n = 25$)			.69
Only assignments rated as fully or partially quantifiable ($n = 23$)			.68
Unquantifiable assignments ($n = 2$)			N/A

Table 3

Frequency of Code Occurrence and Percentage of Total Codes for All Codes on a Nominal or Ordinal Scale

	Frequency	Percentage
Assignment type (<i>n</i> =82)		
One time activity	48	58.5%
Repeated activity	27	32.9%
PRN assignment	2	2.4%
Self-monitoring	3	3.7%
Abstinence/limitation	2	2.4%
Realm of functioning (<i>n</i> =82)		
Non romantic social	8	9.8%
Family/romantic/significant other	8	9.8%
Job/school	15	18.3%
Exercise	7	8.5%
Other health behavior	14	17.1%
Household/logistical	20	24.4%
Other	10	12.2%
Quantifiability (<i>n</i> =82)		
Completely	15	18.3%
Partially quantifiable	56	68.3%
Unquantifiable	11	13.4%
Categorical completion (<i>n</i> =73)		
Fully	34	46.6%
Partially	13	17.8%
Made an attempt or effort to start assignment	6	8.2%
Made no effort to begin assignment	20	27.4%

Appendix

Observer-Based BA Homework Completion Measure

Assignment characterization

1. What type of assignment is it?
 - a. Activity = Patient asked to engage in behavior a single time
 - b. Repeated activity = Patient asked to engage in the same activity multiple times (i.e., across several sittings) between sessions
 - c. PRN assignment = Patient is asked to engage in activity only if another event occurs (e.g., if spouse does something the patient does not like, patient is asked to calmly bring up the issue)
 - d. Self-monitoring = Patient asked to monitor and record behavior or emotions occurring between sessions
 - e. Abstinence/limitation assignment = Patient asked NOT to engage in behavior (e.g., Patient is asked not to drink alcohol) or to engage in a behavior only a limited number of times (e.g., Limit social outings with friends to 2 times/week)
2. What realm of functioning was the assignment in?
 - a. Nonromantic social
 - b. Family/romantic/significant other social
 - c. Job/school
 - d. Exercise (includes any otherwise unspecified physical activity, e.g., hiking, walk dog)
 - e. Other health behavior (e.g., nutrition, checking blood sugar, going to medical appointments, also includes relaxation, breathing, and meditation exercises)
 - f. Household/logistical (e.g., pay bills, clean house, make schedule of family activities, make budget)
 - g. Other, too general
3. How quantifiable is the assignment?
 - a. Completely = Stated clearly in time, distance, amount, or frequency
 - b. Partially quantifiable = Clear when assignment is done, but would be difficult to rate percentage completed
 - c. Unquantifiable = Unclear what behaviors completion would entail

Assignment completion

4. Which best characterizes the patient's completion
 - a. Fully completed assignment
 - b. Partially completed assignment
 - c. Made an attempt or effort to begin assignment
 - d. Made no effort to begin assignment
 5. What percentage of the assignment did the patient complete?

-