The therapeutic alliance in a naturalistic psychiatric setting: Temporal relations with depressive symptom change

Christian A. Webb [a],[*], Courtney Beard [b], Randy P. Auerbach [a], Eliza Menninger [b], Thröstur Björgvinsson [b]

[a] *Center for Depression, Anxiety & Stress Research, Harvard Medical School, McLean Hospital, Department of Psychiatry, 115 Mill Street, Belmont, MA 02478, USA*
[b] *Behavioral Health Partial Hospitalization Program, McLean Hospital, 115 Mill Street, Belmont, MA 02478, USA*

A B S T R A C T

Objective: Numerous studies have reported associations between the therapeutic alliance and depressive symptom improvement in outpatient samples. However, little is known regarding the temporal relationship between the alliance and symptom change among relatively severely depressed patients receiving treatment in naturalistic, psychiatric hospital settings.
Method: Adult patients with major depression ($n = 103$) receiving combined cognitive behavioral therapy and pharmacological treatment at a psychiatric hospital completed repeated assessments of the therapeutic alliance and depressive symptoms, as well as a pretreatment assessment of their expectation of symptom improvement.
Results: Results indicated that the alliance and treatment outcome expectancies significantly predicted subsequent depressive symptom change. However, in a model in which prior symptom change and treatment outcome expectancies were statistically controlled, the alliance-outcome association was rendered nonsignificant. The alliance was significantly associated with prior symptom improvement.
Conclusions: Findings highlight the importance of controlling for plausible third variable and temporal confounds to minimize biased estimates of alliance-outcome associations in future studies. Overall, results were more consistent with the alliance being a consequence, rather than a cause, of symptom change. Finally, findings contribute to a growing body of evidence supporting the role of treatment outcome expectancies in predicting symptom improvement, even within our relatively severely depressed sample.

© 2014 Elsevier Ltd. All rights reserved.

Perhaps no variable has received more attention in the psychotherapy literature than the therapeutic alliance. The most commonly cited definition of the alliance was first articulated by Bordin (1979), who argued that the construct consists of three components: 1) the *bond* between therapist and patient, 2) therapist-patient agreement on the *goals* of treatment, and 3) therapist-patient agreement on the *tasks* of treatment. Indeed, the most commonly used measure of the alliance, the Working Alliance Inventory (WAI; Horvath, Del Re, Flückiger, & Symonds, 2011; Horvath & Greenberg, 1986, 1989), consists of three corresponding subscales designed to assess these components (i.e., Bond, Goals and Tasks subscales).

Several meta-analytic reviews of the alliance-outcome literature have been published over the years (Horvath et al., 2011; Horvath & Symonds, 1991; Martin, Garske, & Davis, 2000). Most recently, Horvath et al. (2011) reported a mean alliance-outcome correlation of .28, indicating that, when averaging across studies, higher alliance ratings are associated with greater symptom improvement. Although intriguing, it is important to note that the vast majority of alliance studies do not control for temporal confounds (Barber, 2009; Webb et al., 2011). In other words, most studies are not predicting subsequent symptom change. More specifically, in the typical alliance study, the alliance is assessed in the midst of treatment and correlated with symptom change from the beginning to the end of treatment. Within such a design, a significant alliance-outcome correlation may be due, at least in part, to the influence of prior symptom change on the alliance. When only considering those studies that have statistically controlled for temporal confounds, alliance-outcome associations are mixed, with

* Corresponding author. Tel.: +1 617 855 4429.
 E-mail address: cwebb@mclean.harvard.edu (C.A. Webb).

some studies reporting that the alliance significantly predicts subsequent symptom change (e.g., Barber, Connolly, Crits-Christoph, Gladis, & Siqueland, 2000; Crits-Christoph, Gibbons, Hamilton, Ring-Kurtz, & Gallop, 2011; De Bolle, Johnson, & De Fruyt, 2010; Falkenström, Granström, & Holmqvist, 2013; Klein et al., 2003; Webb et al., 2011; Zilcha-Mano, Dinger, McCarthy, & Barber, 2013) and others failing to find such an association (e.g., DeRubeis & Feeley, 1990; Feeley, DeRubeis, & Gelfand, 1999; Puschner, Wolf, & Kraft, 2008; Strunk, Brotman, & DeRubeis, 2010; Strunk, Cooper, Ryan, DeRubeis, & Hollon, 2012).

In addition, in the bulk of alliance-outcome studies, ratings of the alliance are based on a single, or just a few, sessions (typically only assessed in the early phase of treatment; see Horvath et al., 2011). Studies using such designs implicitly assume that ratings of the alliance at one, or a few, sessions adequately represent the state of the alliance throughout treatment. A single-session "snapshot" may be sufficient to accurately capture the strength of the alliance. However, if the alliance is relatively unstable over the course of a given treatment, ratings based on only one or a few early sessions would yield unreliable estimates and, consequently, would likely result in relatively weak alliance-outcome associations. In addition, within these studies assessing the alliance at one or a few timepoints, alliance ratings are typically correlated with symptom change over the entire course of treatment, which may fail to capture the shorter-term impact of alliance on symptom improvement. Repeated alliance and symptom assessments over the full course of treatment would allow for a more comprehensive and fine-grained, as well as statistically powerful, test of alliance-outcome associations.

Third, the vast majority of alliance research is conducted within the context of outpatient settings or in carefully controlled clinical trials. Despite the vast body of alliance research published to date, we know surprisingly little about the extent to which the alliance predicts depressive symptom improvement among more severely depressed patients receiving treatment in naturalistic, "real-world" psychiatric settings. Research examining predictors and processes of depressive symptom change in such real-world treatment contexts, in which patients are not carefully selected based on inclusion/exclusion criteria inherent to clinical trials, are critical to informing our understanding of the mechanisms that account for symptom improvement in these naturalistic settings and to compliment data derived from trials.

The goal of the present study was to test the association between the alliance and symptom improvement in a sample of depressed patients while addressing the above mentioned gaps and limitations of prior alliance research. Specifically, we examined the association between the alliance and symptom change in 1) a naturalistic psychiatric setting treating severely depressed patients, while 2) statistically controlling for temporal confounds and 3) assessing both the alliance and depressive symptoms at multiple timepoints throughout treatment. In addition to controlling for temporal confounds, we also wanted to control for plausible third variable confounds of alliance-outcome associations. One can speculate about a number of possibly relevant third variable confounds. However, as others have highlighted, prior symptom change may be one particularly important variable for which to control in alliance-outcome research (Barber et al., 2000; Strunk et al., 2012). Namely, insofar as prior symptom change predicts both subsequent symptom change and alliance scores, it may represent an important third variable for which to control. Indeed, Strunk and colleagues found that the alliance significantly predicted subsequent symptom change in a sample of depressed outpatients. However, in a model in which prior symptom change was statistically covaried, the alliance-outcome association was no longer significant.

Similarly, to the extent that patient expectations of symptom improvement (i.e., *treatment outcome expectancies*) predict both stronger alliances and better treatment outcomes, they may also serve as an important third variable for which to statistically control. As stated by de la Fuente-Fernandez et al. (2001), "the simple act of receiving any treatment (active or not) may, in itself, be efficacious because of expectation of benefit" (p. 1164). Indeed, placebo processes, including the role of treatment outcome expectancies, have received an increased amount of attention in the depression literature in recent years (e.g., Fournier et al., 2010; Kirsch, 2010). Prior research has found that relatively more optimistic treatment outcome expectancies predict greater symptom improvement in depression treatment, including cognitive behavioral therapy (CBT; e.g., Meyer et al., 2002; Webb, Kertz, Bigda-Peyton, & Björgvinsson, 2013). Treatment outcome expectancies have also been shown to be positively correlated with alliance ratings (Constantino, Arnow, Blasey, & Agras, 2005; Joyce et al., 2003; Meyer et al., 2002). Thus, both prior symptom change and pretreatment expectancies may represent two important variables worth statistically controlling in models testing alliance-outcome associations.

As is typically the case in naturalistic psychiatric settings representing higher levels of patient care than outpatient treatment (e.g., inpatient, residential, partial hospitalization units), the psychiatric unit from which the current sample was drawn involved treatment from a multidisciplinary team — including psychologists, psychiatrists, case managers, social workers, occupational therapists and psychiatric nurses — providing group and individual therapy, as well as pharmacological treatment (see *Participants and Treatment Setting* below for details). In contrast, most prior research testing the association between alliance and treatment outcome has been based on individual, one-on-one psychotherapy in outpatient settings. Accordingly, given that patients received their treatment from a psychiatric team rather than a single individual therapist, the alliance with the treatment team as a whole was assessed. Although results from such research may not generalize to traditional outpatient settings, at the same time, the bulk of the alliance-outcome literature to date may not generalize to more acute settings (inpatient, residential, partial hospitalization units), which represent highly utilized — yet understudied — levels of psychiatric care.

We hypothesize that the alliance will significantly predict subsequent depressive symptom change in our sample (*Hypothesis 1a*). However, after controlling for prior symptom change and treatment outcome expectancies, the association between the alliance and subsequent symptom change will no longer be significant (*Hypothesis 1b*). In addition, we expect that the alliance will be significantly positively correlated with prior symptom improvement (*Hypothesis 2*). Finally, informed by prior research, we hypothesize that patient treatment outcome expectancies, assessed pretreatment, will predict greater symptom change (*Hypothesis 3*).

Method

Participants and treatment setting

Participants were patients receiving treatment at the Behavioral Health Partial (BHP) Hospital Program, a partial hospitalization unit at McLean Hospital (Belmont, MA), a Harvard Medical School teaching hospital. To be included in the present study, patients had to be admitted to the BHP and complete the assessment battery described below. Inclusion criteria were that patients met criteria for a current, diagnosis of Major Depressive Disorder, excluding Bipolar Disorder (i.e., current or past Manic/Hypomanic episode), or a current or past Psychotic Disorder. A total of 103 patients (ages

18—68, $M = 36.02$, $SD = 13.71$; 64% females) met these criteria during the study period (July, 2012 to March, 2013) and provided written informed consent for their clinical data to be used for research studies.

Previous episodes of depression were common in our sample, with a mean of 6.42 ($SD = 9.81$) reported previous episodes. The pretreatment mean CES-D-10 depression score for the sample ($M = 20.34$) was double the suggested clinical cutoff of 10 (Andresen, Malmgren, Carter, & Patrick, 1994). Moreover, diagnostic comorbidity was also common, and current diagnoses at the time of intake were as follows: 37.9% ($n = 39$) of our sample met criteria for concurrent Generalized Anxiety Disorder, 23.3% ($n = 24$) for Social Anxiety Disorder, 14.6% ($n = 15$) for Alcohol Abuse/Dependence, 10.7% ($n = 11$) for Obsessive Compulsive Disorder, and 7.8% ($n = 8$) for Post-Traumatic Stress Disorder. Sixty-six percent ($n = 68$) of the current sample met criteria for more than one current DSM-IV disorder, with an average of 2.1 diagnoses ($SD = 1.1$).

The BHP delivers CBT (both group and individual formats) and pharmacological treatment to patients suffering from a wide range of psychiatric disorders (principally mood, anxiety, personality, and psychotic disorders). The program focuses on the acquisition of cognitive behavioral skills and uses a flexible CBT approach, which is adapted to the unique challenges faced in a naturalistic partial hospital setting (see Neuhaus, 2006). Individual treatment plans are constructed for each patient by clinical team managers who conduct intake assessments, develop an initial conceptualization, and oversee all aspects of treatment. Treatment consists primarily of group CBT provided by psychologists, social workers, occupational therapists, postdoctoral and graduate level psychology trainees, and mental health counselors. Patients attend five 50-minute groups each day, five days per week (Monday—Friday). Of these, one group per day focuses on behavioral activation, based on a protocol adapted from Martell, Dimidjian, and Herman-Dunn (2010). A second group is focused on identifying and challenging negative automatic thoughts and guided by a protocol adapted from Beck, Rush, Shaw, and Emery (1979). The remaining group content includes modules on psychoeduction, self-monitoring, mindfulness, and interpersonal skills, adapted from other empirically supported CBT manuals (Beck, Emery, & Greenberg, 1985; Linehan, 1993). To maintain treatment fidelity, groups utilize treatment protocols designed for the program, derived from established treatment manuals. Group leaders are rated for adherence to the protocols twice per year by trained research assistants. Inter-rater reliability is excellent ($r = .99$), and group leaders, on average, address 83% of protocol components during each group (Garner, Stein, Beard, & Bjorgvinsson, 2014). In addition to group therapy and medication consults with a psychiatrist, patients also receive two to three weekly individual CBT sessions from graduate-level psychologists to reinforce and tailor material learned in groups to the patient's unique needs. (For additional details of the BHP, see Beard & Björgvinsson, 2013).

Measures

Mini International Neuropsychiatric Interview (MINI; Sheehan et al., 1998). The MINI is a structured interview assessing for DSM-IV Axis I symptoms (e.g., mood, anxiety, substance abuse, psychosis). Each MINI diagnostic module consists of a series of screening items followed by questions about specific symptomatology. The MINI has strong reliability and validity in relation to the Structured Clinical Interview for DSM-IV (SCID-IV), with inter-rater reliabilities ranging from kappas of .89—1.0 (Sheehan et al., 1998). For the partial hospital patients, inter-rater reliability between the MINI and the program psychiatrists is .69 for MDD and

.75 for Bipolar Disorder-Depressed (Kertz, Bigda-Peyton, Rosmarin, & Björgvinsson, 2012). The MINI was administered by doctoral practicum students and interns in clinical psychology who received weekly supervision by a postdoctoral psychology fellow. Training included reviewing administration manuals and completing mock interviews. All clinicians were required to pass a final training interview with their supervisor before administering MINIs for the program.

Center for the Epidemiological Studies of Depression-10 (CES-D-10; Andersen et al., 1994). The CESD-10 is a widely used, brief patient-report instrument for measuring symptoms of depression. Items assess for symptoms of depression (e.g., "I felt depressed") and response anchors range temporally from 0 = rarely or none of the time to 3 = most or all of the time. The CESD-10 has been shown to have high internal consistency in similar psychiatric samples ($\alpha = .87$; Webb et al., 2013). Given that patients completed the CESD daily, we modified the instructions for the treatment assessments and asked patients to rate the frequency of symptoms over the past 24 h. We altered the 4-point Likert scale accordingly by removing the number of days from each anchor.

Working Alliance Inventory — short version (WAI-S; Tichenor & Hill, 1989; Tracey & Kokotovic, 1989). The WAI-S is a 12-item patient-rated measure of the quality of the therapeutic alliance. Items are rated on a 7-point scale (1 = never to 7 = always). As discussed above, in contrast to conventional outpatient, 1-on-1 therapy, treatment was delivered by a treatment team in an intensive behavioral health partial hospital program at McLean hospital. The patient's treatment team included a psychiatrist, an individual CBT therapist, multiple group therapists, and a case manager. Thus, the wording of WAI items was adapted to reflect the patient's perception of the alliance with his/her *treatment team* as a whole rather than a single individual therapist (e.g., Item 5 was modified from "I am confident in [*therapist's name*]'s ability to help me." to "I am confident in my treatment team's ability to help me. Rate treatment team overall.").

Credibility/Expectancy Questionnaire (CEQ; Devilly & Borkovec, 2000). The CEQ is a 6-item measure divided into two subscales assessing treatment credibility (CEQ-Credibility) and expectancy (CEQ-Expectancy) respectively, and it has demonstrated adequate reliability and validity in previous research (Devilly & Borkovec, 2000). Given that the focus of the current study was on expectancy, we used the CEQ-Expectancy subscale of this measure. Cronbach's alpha was high in the sample ($\alpha = .89$).

Procedure

Patients completed a program orientation on the first day of their treatment. During the orientation, patients were informed that they would complete daily computerized questionnaires to assess their symptoms and functioning, and that this information would be used by the treatment team. They were also given the opportunity to consent for their clinical data to be used for research purposes. Only patients who provided informed written consent are included in the current report. The local Institutional Review Board at McLean Hospital approved all study procedures.

The average duration of treatment for these patients was 11.80 ($SD = 3.81$; Range 2—26) days. Patients completed the CEQ on their first day of treatment (day 1) and the MINI on day 2. The CES-D was administered at day 1, day 2, day 4, day 7 and discharge, for a total of 5 waves of data (henceforth referred to as Time 1—5). The WAI was administered at the same timepoints but the measure was excluded from the Time 1 assessment, as patients did not yet have any interactions with their treatment team to rate the alliance at intake. Given that the study was conducted in a naturalistic clinical (i.e., psychiatric hospital) rather than research setting, some

patients did not complete all of the self-report and diagnostic measures at one or more of the assessment waves for a variety of reasons (e.g., conflicting appointments for individual or group therapy, clinical crises resulting in admission to clinical evaluation center or inpatient unit). Given the abovementioned range in treatment stays, a number of patients did not have time to complete multiple midtreatment assessments (e.g., not completing day 7 assessment). Overall, 86% (89/103) of our sample had at least 4 waves of data, to allow us to test the association between the alliance and depressive symptom change over the course of treatment (39 patients had 5 waves of complete data, 50 had 4 waves, 13 had 3 waves, and 1 patient had only 2 waves of data).[1]

Study data were collected and managed using the REDCap (Research Electronic Data Capture) application hosted at McLean Hospital. REDCap is a web-based application designed to facilitate and securely streamline data capture for research studies (Harris et al., 2009).

Analytic strategy

To model the association between alliance (WAI) scores and depressive symptom (CES-D) change over time, and similar to Strunk et al. (2012), a repeated measures regression was conducted using SAS (9.2) mixed procedure and maximum likelihood estimation. First, to test the association between alliance and *subsequent* symptom change, a vector of lagged CES-D scores for each patient served as the dependent variable (i.e., CES-D at Time 3 through 5), with CESD-D scores at the previous timepoint (Time T-1) entered as covariates. A vector of WAI scores was entered as our predictor variable (i.e., WAI at Time 2–4). That is, the latter model uses repeated assessments to statistically estimate the relation between the WAI (Time 2–4) and CES-D scores at the next timepoint (Time $T+1$), adjusting for CES-D scores at the same timepoint as the WAI assessment (Time 2–4). For the latter model, a negative t value indicates that higher scores on the WAI predicted lower subsequent CES-D depressive symptoms.

Next, and in line with prior research (e.g., Strunk et al., 2012; Webb et al., 2011), to test whether *prior* symptom change predicted the strength of the alliance, a vector of lagged WAI scores for each patient served as the dependent variable (Time 2–5), and a residualized prior CES-D change score (i.e., CES-D at Time T, adjusting for CES-D at Time 1) was entered as a covariate. A positive t value indicates a positive association between prior symptom improvement and subsequent alliance scores.[2]

Within studies including repeated assessments obtained from the same individuals over time (and consequently within-subject observations, and residuals, over time are likely to be correlated), commonly used covariance structures include first-order autoregressive, heterogeneous autoregressive, banded Toeplitz, and compound symmetry. In order to select between these different covariance structure options, we fitted models using each structure

and chose the best fit based on Akaike's information criterion, Schwarz's Bayesian criterion, and -2 log likelihood. For analyses predicting subsequent symptom change, the best fit was banded Toeplitz; whereas for analyses predicting alliance scores, the best fit was heterogeneous autoregressive.

Results

Means, standard deviations, and correlations for CES-D and WAI variables at each timepoint are listed in Table 1. As expected, CES-D and WAI scores were significantly negatively correlated with one another, and scores on the same measure over time were positively correlated. Greater (i.e., more optimistic) treatment outcome expectancies (CEQ-Expectancy) were associated with significantly lower CES-D scores at each timepoint and significantly positively correlated with WAI scores at all timepoints, with the exception of Time 4.

Conventional analyses: alliance-outcome associations by timepoint

As noted above, the bulk of prior alliance-outcome research has tested the association between the alliance at a given session and symptom change (see recent meta-analytic review by Horvath et al., 2011). However, for the primary analyses in the current study, we chose to take advantage of the repeated assessment structure of our data and run a repeated measures regression testing the association between alliance and subsequent (or prior) symptom change over time in a single statistical step. Nevertheless, for ease of comparison with prior research, we present here analyses examining the association, for each timepoint separately, between (1) WAI scores at time T and subsequent depressive symptom change (i.e., predicting CES-D at Time $T+1$, adjusting for both CES-D at time T and a residualized prior symptom change score) and (2) WAI scores at Time T and residualized prior change scores (i.e., CES-D at Time T, adjusting for CES-D at Time 1). See Fig. 1 below for a graphical summary of our results at each timepoint. The WAI did not significantly predict subsequent CES-D symptom improvement at any timepoint: WAI assessed at Time 2 (partial $r = -.17$, $p = .101$), Time 3 (partial $r = -.19$, $p = .136$) and Time 4 (partial $r = -.07$, $p = .681$). In contrast, higher WAI scores were significantly associated with greater prior CES-D symptom improvement at all timepoints: WAI assessed at Time 2 ($r = .24$, $p = .016$), Time 3 ($r = .25$, $p = .016$), Time 4 ($r = .38$, $p = .001$), and Time 5 ($r = .40$, $p = .002$).

Repeated measures regression

Prediction of subsequent symptom change

Next, we conducted a repeated measures regression to test the association between the alliance and subsequent symptom change over time in a single statistical step. As shown in Table 2 (top panel), higher alliance scores significantly predicted greater subsequent symptom improvement. However, after controlling for prior depressive symptom change and treatment outcome expectancies, the alliance-outcome association was reduced to a nonsignificant trend (see Table 2: Panel 2). In this model, relatively higher (i.e., more optimistic) treatment outcome expectations were associated with greater symptom improvement.

To examine whether certain components of the alliance might be more strongly associated with symptom change, we conducted a series of repeated measures regression analyses testing the association between the three constituent subscales of the WAI (Bond, Task and Goals) and subsequent depressive symptom change (again controlling for prior symptom change and expectancies). Only the Task subscale emerged as a significant predictor of subsequent

---

[1] To statistically test whether duration of treatment (i.e., total number of partial hospital treatment days) moderates either (1) the relation between the alliance and subsequent symptom change and/or (2) the association between residualized prior symptom change and subsequent alliance scores, we added a treatment duration interaction term to the models tested above following the recommendation of Aiken and West (1991); i.e., prior to computing cross-products for inclusion in tests of interaction effects [predicting, depending on the analysis, either CES-D subsequent symptom change or WAI scores], relevant variables were mean-centered. Duration of treatment did not moderate any of these associations (all $ts < .95$ and $ps > .35$ for interaction terms).

[2] Several outliers were identified and excluded from our analyses. Specifically, one outlier for CES-D at Time 1 ($z = -3.61$), the WAI at Time 4 ($z = -3.09$), as well as the WAI Goals subscale at Time 2 ($z = -3.06$) and WAI Bond subscale at Time 5 ($z = -3.61$) were detected and deleted.

Table 1
Means, standard deviations, and correlations for all variables.

Variable	M	SD	2	3	4	5	6	7	8	9	10
1.CES-D_T1	20.34	5.24	.44**	.41**	.40**	.48**	−.22*	−.28**	−.10	−.30*	−.23*
2. CES-D_T2	14.76	5.27	–	.39**	.36**	.37**	−.29**	−.10	−.08	−.14	−.23*
3. CES-D_T3	13.01	6.11		–	.71**	.55**	−.29**	−.37*	−.31**	−.36**	−.31**
4. CES-D_T4	11.37	6.21			–	.63**	−.21	−.43**	−.37**	−.60**	−.41**
5. CES-D_T5	10.65	6.13				–	−.34**	−.41**	−.46**	−.49**	−.27*
6. WAI_T2	5.01	1.17					–	.57**	.25*	.39**	.31**
7. WAI_T3	5.54	0.94						–	.35**	.72**	.30**
8. WAI_T4	5.75	1.07							–	.86**	.17
9. WAI_T5	6.00	0.90								–	.42**
10. CEQ-Exp	15.85	5.50									–

Note: CES-D = Center for the Epidemiological Studies of Depression-10; WAI = Working Alliance Inventory; CEQ-Exp = Credibility/Expectancy Questionnaire − Expectancy subscale; T1−T5 = Time 1−5 assessments; *p < .05. **p < .01.

symptom improvement ($b = −.17$; SE = .08; $t = −2.07$; $p = .041$); whereas the Bond subscale was associated with greater symptom change at the level of a nonsignificant trend ($b = −.13$; SE = .08; $t = −1.68$; $p = .096$; and for the Goals subscale, $b = −.11$; SE = .08; $t = −1.37$; $p = .173$).

Prediction of the alliance from prior symptom change

The alliance was significantly associated with prior symptom change, in the direction of greater symptom improvement predicting higher alliance scores (See Table 2; Panel 3). This association remained significant after controlling for treatment outcome expectancies (See Table 2; Panel 4). Relatively higher (i.e., more optimistic) treatment outcome expectations were associated with higher alliance ratings. Each of the constituent alliance subscales was significantly associated with greater prior symptom change (Task subscale, $b = .21$; SE = .05; $t = 4.30$; $p < .001$; Goals subscale, $b = .19$; SE = .05; $t = 3.96$; $p < .001$; Bond subscale, $b = .17$; SE = .05; $t = 3.57$; $p < .001$).

Discussion

The present study examined the association between the therapeutic alliance and depressive symptom improvement while addressing several limitations of prior alliance research. Specifically, in contrast to the vast majority of prior alliance studies, we tested alliance-outcome associations within 1) the context of a naturalistic psychiatric setting treating severely depressed patients, while 2) assessing the alliance at multiple timepoints throughout treatment, as well as 3) statistically controlling for temporal confounds and 4) two plausible third variable confounds.

Results indicated that the alliance was a significant predictor of subsequent depressive symptom improvement. However, after statistically controlling for prior symptom change and patient treatment outcome expectancies, the alliance-outcome association was no longer statistically significant. These findings parallel those of Strunk et al. (2012) who found a similar pattern of findings in the context of a randomized clinical trial (RCT) of combination CBT and antidepressant medication for depressed outpatients. In contrast, the current sample was not drawn from an RCT with carefully selected patients, but rather from a naturalistic, "real-world" psychiatric hospital unit treating relatively severely depressed patients (who were also receiving combined CBT and pharmacotherapy). Overall, the current findings suggest that future investigations testing alliance-outcome associations should control for prior symptom change and treatment outcome expectancies, as well as other plausible third variable confounds which, if uncontrolled, may bias estimates of alliance-outcome relations.

It should be noted that although total alliance scores were not significantly associated with depressive symptom improvement after statistically controlling for the included third variables, when examining the three constituent subscales of the WAI, the Task (but not the Bond or Goal) subscale did emerge as a significant predictor of depressive symptom change. These findings may highlight the importance − in CBT − of agreement on the concrete tasks of treatment in contributing to depressive symptom improvement (For similar findings see Strunk et al., 2012 and Webb et al., 2011).
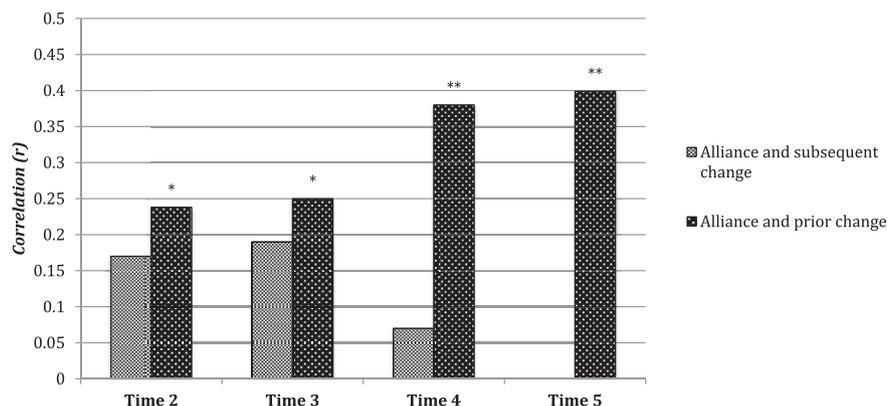

Fig. 1. Two sets of correlations were conducted testing the association between the alliance and both prior (black bars) and subsequent (gray bars) symptom change, at each timepoint separately. Specifically, the first set of bars (gray) represent the correlation between alliance scores and subsequent symptom change, partialing out a residualized prior CES-D change score (i.e., CES-D at Time *t*, adjusting for Time 1). The second set of bars (black) represents the correlation between alliance scores at time *t* and a residualized prior CES-D change score. For ease of comparison, for both sets of bars, positive correlations reflect that higher alliance scores are associated with greater subsequent and prior symptom improvement. **p < .01; *p < .05.

Table 2
Repeated measures regression analyses of the relation between alliance and both prior and subsequent symptom change.

Predictor (s)	Dependent variable	Model	Parameter estimate (b)	Standard Error (SE)	t-value	p-value
Alliance	Subsequent symptom change	1: *Basic Model*	−.08	.03	−2.69	.008
Prior symptom change	Subsequent symptom change	2: *Covarying Prior CES-D*	.59	.17	3.56	.001
Expectancies		*Change & Expectancies*	−.17	.08	−2.14	.035
Alliance			−.05	.03	−1.71	.090
Prior symptom change	Alliance	3: *Basic Model*	.49	.12	4.10	.000
Expectancies	Alliance	4: *Covarying Expectancies*	.60	.17	3.58	.001
Prior symptom change			.45	.12	3.78	.000

Note: Alliance = Working Alliance Inventory; Subsequent symptom change = A vector of lagged CES-D scores (i.e., CES-D at the timepoint subsequent to the alliance assessment), with CES-D scores at the previous timepoint (T-1) entered as a covariate. Prior symptom change = CES-D symptom change prior to alliance assessment. Degrees of freedom (df) for model 1 (df = 106), model 2 (df = 104) and model 3 (df = 222) and model 4 (df = 222). Unstandardized estimates (b), standard errors (SE) and both t-values and p-values are provided. In models 1 and 2, negative t values reflect a predictor variable being related to larger improvements in symptom scores. In models 3 and 4, positive t values reflect a predictor variable being associated with higher alliance scores.

Patient treatment outcome expectancies also predicted greater depressive symptom improvement. In other words, those patients who entered treatment with greater (i.e., more optimistic) expectations of symptom improvement tended to experience relatively more symptom change over the course of treatment. Similarly, greater expectations of symptom improvement also predicted significantly stronger alliance scores at three out of the four assessment timepoints (see Table 1). It is interesting to note that the "mere" expectation of therapeutic benefit predicted symptom change in this relatively severe patient population receiving treatment in a short-term partial hospital setting (i.e., average length of stay of 11.8 days). These findings contribute to a growing body of research highlighting the role of patient expectations of symptom improvement in predicting the extent of symptom change, and suggest that expectancy-outcome effects may extend to highly symptomatic patients (Meyer et al., 2002; Webb et al., 2013). With regards to clinical implications, these results may underscore the importance of clinicians assessing — and ideally fostering optimistic (yet realistic) — expectations of improvement in their depressed patients, many of whom may hold — in line with the cognitive model — relatively pessimistic cognitions about their treatment prognosis. Of course, it must be highlighted that these are observational, non-experimental, findings and thus conclusions about the role of expectancies in driving symptom change must be tempered. Given the observational nature of the study there may have been certain unmeasured third variable confounds (e.g., particular patient characteristics) which influenced our expectancy-outcome results.

To take advantage of the multiple assessment timepoints, we conducted repeated measures regressions in our primary analyses to test the association between the alliance and depressive symptom change over time. However, an examination of the alliance-outcome associations at each assessment timepoint separately may prove informative. Specifically, as illustrated in Fig. 1, at no timepoint did the alliance significantly predict subsequent symptom change, with correlations in the small range (rs −.07 to −.19). In contrast, the alliance was significantly associated with prior symptom change at all timepoints, with correlations ranging from .24 early in treatment to .40 at posttreatment. Taken together, these findings suggest that the alliance may be more of a consequence, than a cause, of depressive symptom change, at least in the context of the present CBT-based partial hospital setting. Interestingly, the most recent meta-analysis by Horvath et al. (2011) found larger alliance-outcome associations within those studies that assessed the alliance later in treatment. Given the fact that the vast majority of these alliance studies do not control for temporal confounds (e.g., instead testing the association between the alliance, assessed in the midst of treatment, and pre-posttreatment symptom change), one can speculate that the

larger alliance-outcome associations obtained in studies assessing the alliance later (relative to earlier) in treatment may reflect, at least in part, the association between the alliance and *prior* symptom improvement. The present findings highlight the need for future studies adequately controlling for temporal confounds, and suggest that the failure to do so may artificially inflate estimates of alliance-outcome associations due to the influence of prior symptom improvement on subsequent alliance ratings. It should also be noted that, in our individual timepoint analyses, we did not find that the alliance significantly predicted subsequent symptom change (although the effects were in the expected direction). In contrast, in the repeated measures regression analyses the alliance was found to be a significant predictor of symptom change (when not controlling for expectancies and prior symptom change). This apparent discrepancy may be due, at least in part, to the greater statistical power achieved in a repeated measures regression including all available timepoints.

Several limitations of the present study should be noted. The study was conducted within the context of a short-term naturalistic behavioral health partial hospital program, with an average length of stay of approximately 12 days. The extent to which our findings generalize to other treatment settings is unclear. However, as noted above, Strunk et al. (2012) did obtain a similar pattern of findings in an outpatient sample. In addition, given that patients received treatment from a multidisciplinary team, rather than from an individual therapist, we used an adapted version of the WAI in order to assess the alliance with the treatment team as a whole. Thus, it is unclear to what extent our findings may generalize to other settings (e.g., individual psychotherapy) and to research on the traditional "one-on-one" therapist-patient alliance construct. Although our results may not generalize to the typical outpatient setting, at the same time, the literature to date may not generalize to more acute settings (inpatient, residential, partial hospital units), which represent important and highly utilized — yet understudied- levels of care that should not be ignored. In addition, the alliance was assessed via patient report, which is the most common method of assessing the alliance (Horvath et al., 2011). Thus, it is not clear to what extent our findings would generalize to other assessment methods (e.g., therapist report or observational coding via trained raters). Moreover, the importance of taking the hierarchical structure of treatment data (i.e., patients nested within therapists) into account in statistical models has received increased attention in the psychotherapy literature (see Baldwin, Wampold, & Imel, 2007; Crits-Christoph et al., 2011). However, given the complexities of patient care in our naturalistic setting (e.g., patients were each treated by a variety of different mental health providers and attended a range of different groups during their stay) the creation of nested models were not deemed feasible within the current study.

Future directions

Despite the vast body of alliance-outcome research published to date, substantial debate persists regarding the role of the therapeutic alliance as a cause, versus consequence, of depressive symptom change. Study designs in alliance research have historically been limited by critical temporal confounds. More recent research has controlled for temporal, and several plausible third variable, confounds. However, given the observational nature of previous alliance studies, including the current investigation, unmeasured third variables may in theory have contributed to observed alliance-outcome associations. Ultimately, well-designed experimental studies are needed in which researchers manipulate the strength of the alliance (or particular components of the alliance, e.g., Tasks), while attempting to hold all other process variables constant and randomly assigning patients to alliance conditions. Great methodological care and creativity would undoubtedly be required in such future studies given the challenges of manipulating a dyadic variable such as the alliance (as opposed to manipulating therapist delivery of particular therapy techniques). Nevertheless, such studies are sorely needed to maximize internal validity and strengthen causal inferences regarding the role of the alliance (or particular components of the construct) in driving symptom change depression treatment. More broadly, such well-designed studies testing the associations between theoretically important therapy process variables (not only the alliance) and treatment outcome are needed to generate meaningful clinical recommendations to ultimately improve the efficacy of our current treatments for depression.

Acknowledgments

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Andresen, E. M., Malmgren, J. A., Carter, W. B., & Patrick, D. L. (1994). Screening for depression in well older adults: evaluation of a short form of the CES-D (center for epidemiologic studies depression scale). *American Journal of Preventive Medicine, 10*, 77–84.

Baldwin, S. A., Wampold, B. E., & Imel, Z. E. (2007). Untangling the alliance–outcome correlation: exploring the relative importance of therapist and patient variability in the alliance. *Journal of Consulting and Clinical Psychology, 75*, 842–852.

Barber, J. P. (2009). Towards a working through of some core conflicts in psychotherapy research. *Psychotherapy Research, 19*, 1–12.

Barber, J. P., Connolly, M. B., Crits-Christoph, P., Gladis, L., & Siqueland, L. (2000). Alliance predicts patients' outcome beyond in-treatment change in symptoms. *Journal of Consulting and Clinical Psychology, 68*, 1027–1032.

Beard, C., & Björgvinsson, T. (2013). Commentary on psychological vulnerability: an integrative approach. *Journal of Psychotherapy Integration, 23*(3), 281–283.

Beck, A. T., Emery, G., & Greenberg, R. L. (1985). *Anxiety disorders and phobias: A cognitive perspective*. New York, NY: Basic Books.

Beck, A. T., Rush, J. A., Sahw, B. F., & Emery, G. (1979). *Cognitive therapy for depression*. New York, NY: Guilford Press.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research, and Practice, 16*, 252–260.

Constantino, M. J., Arnow, B. A., Blasey, C., & Agras, W. S. (2005). The association between patient characteristics and the therapeutic alliance in cognitive-behavioral and interpersonal therapy for bulimia nervosa. *Journal of Consulting and Clinical Psychology, 73*(2), 203.

Crits-Christoph, P., Gibbons, M. B. C., Hamilton, J., Ring-Kurtz, S., & Gallop, R. (2011). The dependability of alliance assessments: the alliance–outcome correlation is larger than you might think. *Journal of Consulting and Clinical Psychology, 79*(3), 267.

De Bolle, M., Johnson, J. G., & De Fruyt, F. (2010). Patient and clinician perceptions of therapeutic alliance as predictors of improvement in depression. *Psychotherapy and Psychosomatics, 79*, 378–385.

DeRubeis, R. J., & Feeley, M. (1990). Determinants of change in cognitive therapy for depression. *Cognitive Therapy and Research, 14*, 469–482.

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

Falkenström, F., Granström, F., & Holmqvist, R. (2013). Therapeutic alliance predicts symptomatic improvement session by session. *Journal of Counseling Psychology, 60*(3), 317–328. http://dx.doi.org/10.1037/a0032258.

Feeley, M., DeRubeis, R. J., & Gelfand, L. (1999). The temporal relation of adherence and alliance to symptom change in cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 67*, 578–582.

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

de la Fuente-Fernandez, R., Ruth, T. J., Sossi, V., Schulzer, M., Calne, D. B., & Stoessl, A. J. (2001). Expectation and dopamine release: mechanism of the placebo effect in Parkinson's disease. *Science, 293*, 1164–1166.

Garner, L., Stein, A., Beard, C., & Bjorgvinsson, T. (2014, January). *Assessing treatment fidelity in a naturalistic treatment setting*. Poster presented at the 5th annual meeting of McLean Hospital Research Day, Belmont, MA.

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, J., & Conde, D. (2009). Research electronic data capture (REDCap) — a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics, 42*(2), 377–381.

Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy, 48*(1), 9.

Horvath, A. O., & Greenberg, L. S. (1986). The development of the working alliance inventory. In L. S. Greenberg, & W. M. Pinsoff (Eds.), *The psychotherapeutic process: A research handbook* (pp. 529–556). New York: Guilford Press.

Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the working alliance inventory. *Journal of Counseling Psychology, 36*, 223–233.

Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: a meta-analysis. *Journal of Counseling Psychology, 38*, 139–149.

Joyce, A. S., Ogrodniczuk, J. S., Piper, W. E., & McCallum, M. (2003). The alliance as mediator of expectancy effects in short-term individual therapy. *Journal of Consulting and Clinical Psychology, 71*(4), 672–679. http://dx.doi.org/10.1037/0022-006X.71.4.672.

Kertz, S. J., Bigda-Peyton, J. S., Rosmarin, D. H., & Björgvinsson, T. (2012). The importance of worry across diagnostic presentations: prevalence, severity and associated symptoms in a partial hospital setting. *Journal of Anxiety Disorders, 26*(1), 126–133.

Kirsch, I. (2010). *The Emperor's new drugs: Exploding the antidepressant myth*. London: The Bodley Head.

Klein, D. N., Schwartz, J. E., Santiago, N. J., Vivian, D., Vocisano, C., Castonguay, L. G., et al. (2003). Therapeutic alliance in depression treatment: controlling for prior change and patient characteristics. *Journal of Consulting and Clinical Psychology, 71*, 997–1006.

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. New York, NY: Guilford Press.

Martell, C. R., Dimidjian, S., & Herman-Dunn, R. (2010). *Behavioral activation for depression: A clinician's guide*. New York, NY: Guilford Press.

Martin, D. J., Garske, J. P., & Davis, K. M. (2000). Relation of the therapeutic alliance with outcome and other variables: a meta analytic review. *Journal of Clinical and Consulting Psychology, 68*, 438–450.

Meyer, B., Pilkonis, P. A., Krupnick, J. L., Egan, M. K., Simmens, S. J., & Sotsky, S. M. (2002). Treatment expectancies, patient alliance and outcome: further analyses from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology, 70*(4), 1051–1055. http://dx.doi.org/10.1037/0022-006X.70.4.1051.

Neuhaus, E. C. (2006). Fixed Values and a Flexible Partial Hospital Program Model. *Harvard Review of Psychiatry, 14*(1), 1–14. http://dx.doi.org/10.1080/10673220500519706.

Puschner, B., Wolf, M., & Kraft, S. (2008). Helping alliance and outcome in psychotherapy: what predicts what in routine outpatient treatment? *Psychotherapy Research, 18*, 167–178. http://dx.doi.org/10.1080/10503300701367984.

Sheehan, D. V., Lecrubier, Y., Sheehan, H. K., Amorim, P., Janavs, J., Weiller, E., et al. (1998). The mini-international neuropsychiatric interview (M.I.N.I.): the development and validation of a structured diagnostic interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry, 59*(Suppl. 20), 22–33.

Tichenor, V., & Hill, C. E. (1989). A comparison of six measures of working alliance. *Psychotherapy, 26*, 195–199.

Tracey, T. J., & Kokotovic, A. M. (1989). Factor structure of the working alliance Inventory. *Psychological Assessment, 1*, 207–210.

Strunk, D. R., Brotman, M. A., & DeRubeis, R. J. (2010). The process ofchange in cognitive therapy for depression: Predictors of early intersessionsymptom gains. *Behaviour Research and Therapy, 48*, 599—606. http://dx.doi.org/10.1016/j.brat.2010.03.011.

Strunk, D. R., Cooper, A. A., Ryan, E. T., DeRubeis, R. J., & Hollon, S. D. (2012). The process of change in cognitive therapy for depression whencombined with antidepressant medication: predictors of early intersessionsymptom gains. *Journal of Consulting and Clinical Psychology, 80*, ,730—738.

Webb, C. A., DeRubeis, R. J., Amsterdam, J. D., Shelton, R. C., Hollon, S. D., & Dimidjian, S. (2011). Two aspects of the therapeutic alliance: differential relations with depressive symptom change. *Journal of Consulting and Clinical Psychology, 79*(3), 279.

Webb, C. A., Kertz, S. J., Bigda-Peyton, J. S., & Björgvinsson, T. (2013). The role of pretreatment outcome expectancies and cognitive—behavioral skills in symptom improvement in an acute psychiatric setting. *Journal of Affective Disorders, 49*(1), 375—382.

Zilcha-Mano, S., Dinger, U., McCarthy, K. S., & Barber, J. P. (2013, November 25). Does alliance predict symptoms throughout treatment, or is it the other way around? *Journal of Consulting and Clinical Psychology.* http://dx.doi.org/10.1037/a0035141. Advance online publication.