



HHS Public Access

Author manuscript

Clin Psychol Sci. Author manuscript; available in PMC 2024 March 08.

Published in final edited form as:

Clin Psychol Sci. 2021 May ; 9(3): 515–523. doi:10.1177/2167702620959352.

Reevaluating the Alliance–Outcome Relationship in the Early Sessions of Cognitive Behavioral Therapy of Depression

Megan L. Whelen,

Samuel T. Murphy,

Daniel R. Strunk

The Ohio State University

Abstract

The therapeutic alliance has been reliably associated with outcome across psychotherapies. We investigated the alliance-outcome relationship in the early sessions of cognitive behavioral therapy of depression using a model that disaggregates within- and between-person variance while estimating the reciprocal relation between variables. We utilized this model in a combined dataset from two studies totaling 191 patients. In our primary model, we found evidence for a predictive within-patient relationship between alliance and symptoms such that symptoms predicted regressed change in alliance and alliance predicted regressed change in symptoms. In a more conservative detrended model, these relationships were not significant. Given that a) most of the variability in alliance scores is between-patient; b) the size of the alliance-outcome relationship is modest; and c) the alliance-outcome relationship is not robust to detrending, our findings suggest the alliance plays at most a small role in improving patient outcomes in cognitive behavioral therapy of depression.

Keywords

Alliance; alliance-outcome; depression; cognitive behavioral therapy; psychotherapy; longitudinal methods

The therapeutic alliance is likely the most extensively studied psychotherapy process variable. Several meta-analyses have aggregated this body of literature to generate estimates of the alliance-outcome relationship (Flückiger, Del Re, Wampold, & Horvath, 2018; Martin, Garske, & Davis, 2000), with evidence suggesting significant heterogeneity in these estimates. However, insofar as many constituent studies have methodological limitations, estimates in the literature may be much less informative than desired. Recently developed analytic approaches to modeling the reciprocal relations in panel data have several advantages in testing potential causal relations. Given the heterogeneity of estimates observed and the view that the role of the alliance may vary in different treatments, it is important to test these models in a variety of samples (Muran & Barber, 2011). To date, only a few studies have examined the alliance and outcome using such a modeling approach.

These studies have tended to examine heterogeneous treatments for diagnostically diverse patient samples. This is particularly important as some research has suggested that the alliance may play a less important role in CBT than in treatment approaches in which the relationship is posited to play a more central role (Zilcha-Mano, Eubanks, & Muran, 2019). In this paper, we report on our effort to model the reciprocal relation of the alliance and outcome in CBT specifically.

Desirable Methodological Features

When examining the relation of the alliance and outcome, there are several important features that are desirable. Ideally, testing the impact of the alliance on outcome would include measurements that precede the symptom improvement these measures may predict (Feeley, DeRubeis, & Gelfand, 1999). If the predictor variable temporally precedes the dependent variable, any such relation is not likely to be accounted for by reverse causality. In some studies, the outcome variable has been the difference between a post-treatment measurement and a measurement at intake, while the alliance variable is measured mid-treatment. In this case, a relation between the alliance and change in symptoms could be accounted for in several ways (i.e., the alliance causing outcome, a reverse-causal relation, or a third variable introducing a spurious relation).

Repeated assessments of the alliance and outcome offer multiple advantages. They offer increased statistical power and, with appropriately spaced assessments, allow for assessing the lag of interest. Although the optimal lag for these relationships is often not known, methodologists recommend the assessment frequency be chosen to capture the lag the researchers hypothesize is operating (Ebner-Priemer & Trull, 2009). In the context of psychotherapy, changes from session to session are likely to provide a reasonably fine-grained characterization of the alliance that correspond to periods over which one might see symptom improvement (but see Eubanks, Muran, & Safran, 2018).

Repeated assessment also allows researchers to disaggregate within- and between-person effects (Wang & Maxwell, 2015). When repeated measures are collected, any association identified may be driven by the between-person or within-person variation. While between-person associations might be due to a variety of stable personal characteristics or third variables, within-person associations are not. While some more recent studies assessing alliance and outcome have taken this approach (e.g., Fisher, Atzil-Slonim, Bar-Kalifa, Rafaeli, & Peri, 2016; Rubel, Rosenbaum, & Lutz, 2017), the vast majority of the alliance-outcome literature is composed of studies that did not disaggregate within- and between-person variability. Such studies leave open the possibility that results are confounded by a stable patient characteristic that may influence both the predictor and the dependent variable (Allison, 2014; Curran & Bauer, 2011; Wang & Maxwell, 2015).

Another desirable feature in research testing the alliance-outcome relationship with repeated measures is to control for a lagged dependent variable. Insofar as there is either a true causal effect of a variable on itself at a later time point or a correlation between error in the measurement of a variable across time, models need to account for a lagged dependent variable (Allison, 1990). Unfortunately, common approaches to including a

lagged dependent variable can be problematic. Allison (2015) has suggested that such an approach should be avoided as it introduces “severe bias” to model estimates. Specifically, the coefficient for the lagged dependent variable is usually too large, and the other coefficients in the model are usually too small (Allison, 2015; Falkenström et al., 2017).

Fortunately, some recent modeling approaches do allow one to include a lagged dependent variable without introducing bias. One such approach that allows for each of the desirable features we have described is Hamaker, Kuiper, and Grasman’s (2015) Random Intercept Cross-Lagged Panel Model (RI-CLPM). The RI-CLPM allows for examination of within-person relationships among repeated measurements of variables. Like a traditional cross-lagged panel model, it allows for modeling the effect of X on Y while also modeling the effect of Y on X. Such a reciprocal model is appropriate for studying the alliance and outcome as experts have suggested that the alliance may contribute to the outcome and that outcome may also contribute to the alliance. In fact, empirical findings of both of these relationships have often been reported (Flückiger et al., 2018; Webb, Beard, Auerbach, Menninger, & Björgvinsson, 2014). The RI-CLPM also parses between-person vs. within-person variation, which allows testing whether a within-person relationship is evident. The RI-CLPM is useful for testing whether and how two variables are related and determining the relative magnitude of each variable’s predictive relation with the other (Hamaker et al., 2015). Given these substantial advantages, we used the RI-CLPM to examine the reciprocal relationship between alliance and symptoms in CBT of depression.

Several recent papers have tested reciprocal alliance-outcome models. The samples used in these papers have often had a diverse range of presenting problems and included a variety of treatments. Three papers have examined these models in samples of patients that varied widely in their presenting problems (including depression, anxiety, substance use disorders, and psychosis) and in the treatments that were used (Falkenström, Granstrom, & Holmqvist, 2014; Falkenström, Kuria, Othieno, & Kumar, 2019; Xu & Tracey, 2015). We are aware of only one study that focused on depression treatments. In that study, Zilcha-Mano, Dinger, McCarthy, and Barber (2014) found support for a reciprocal alliance-outcome relationship in a depressed sample who received either supportive–expressive psychotherapy, clinical management combined with pharmacotherapy, or clinical management combined with placebo. However, it is important to investigate the alliance-outcome relationship in specific treatments and patient populations. Several researchers have suggested that the importance of the alliance may vary across treatments depending on the extent to which the therapist-client relationship is central in each treatment (e.g., Siev, Huppert, & Chambless, 2009). There is also evidence to suggest that the alliance-outcome relationship varies as a function of patient characteristics (Lorenzo-Luaces, DeRubeis, & Webb, 2014; Zilcha-Mano & Errázuriz, 2015). For example, this relationship appears stronger among depressed patients with fewer prior episodes (Lorenzo-Luaces, DeRubeis, & Webb, 2014). In a study that separated within and between patient variation, Zilcha-Mano and Errázuriz (2015) found that the within-patient alliance-outcome relationship was stronger among patients with greater symptom severity in their study of patients participating in psychotherapy in a general mental health clinic in Chile. Such findings suggest it will be important to apply analytic approaches with desirable features to patients participating in different treatments and presenting with different clinical problems.

To our knowledge, this paper is the first to report on examination of a reciprocal alliance-outcome model for CBT of depression. We hypothesized that there would be a significant reciprocal relationship between alliance and symptoms, where alliance significantly predicts subsequent regressed change in symptoms and symptoms significantly predict subsequent regressed change in alliance when these relations are modeled together.

Methods

Participants

We drew from two samples of patients participating in CBT of depression. Both samples were of patients with a primary diagnosis of major depressive disorder (MDD) as assessed by the Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders-IV (First, Spitzer, Gibbon, & Williams, 2002). The first sample was composed of 66 patients (see Adler, Strunk, & Fazio, 2015) and the second sample consisted of 125 patients (see Schmidt, Pfeifer, & Strunk, 2019). For each dataset, one additional participant could not be included as they discontinued after their intake and before participating in any sessions. For both samples, inclusion criteria were current MDD, being 18 years of age or older, and providing informed consent. Exclusion criteria were bipolar I, history of psychosis, presence of a primary diagnosis other than MDD that necessitated treatment other than that being offered, current suicide or selfharm risk precluding outpatient treatment, substance dependence in the past 6 months, or clear indication of secondary gain. The first study also planned to exclude patients with an IQ below 80 (with testing only occurring when clinically indicated), but no patients were excluded on this basis. The second study also excluded patients with bipolar II. Combining these samples yielded a total of 191 patients.

In the combined sample, most participants were Caucasian (82%), with the largest remaining groups being African American (8%) and Asian American (7%), with the modal patient having completed some college (39%). The majority of participants were female (58%) and the mean age of our sample was 33.5 (SD = 13.3, range = 18 – 70). All participants provided consent, and study procedures were approved by our institutional review board.

Therapists

Therapists were nine graduate students (five men and four women) under the supervision of a licensed clinical psychologist (DRS). Therapists were pseudo-randomly assigned to patients, with assignments made based on openings in each therapist's caseload, and the patient's intake assessor was not permitted to serve as the patient's therapist. The study protocol called for twice weekly therapy sessions for the first four weeks, with time between meetings collaboratively established as weekly or bi-weekly after four weeks until the final four weeks of treatment. For the final four weeks of therapy the study protocol called for once-weekly sessions.

Depression severity.—The Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996) is a 21-item self-report measure of depressive symptom severity participants

completed prior to each session. The BDI-II is a widely used measure with excellent psychometric properties (Beck et al., 1996).

Working alliance.—The short form of the Working Alliance Inventory – Client Version (WAI; Horvath & Greenberg, 1989; Tracey & Kokotovic, 1989) consists of 12 items rated on a 7-point Likert-type scale. The WAI shows excellent and robust psychometric properties (Horvath & Greenberg, 1989). Participants filled out this measure following each of their first four sessions.

Analytic strategy

We used a RI-CLPM model implemented in Mplus version 8.3 (Muthén & Muthén, 2017). Although BDI data were available at every session, we only collected WAI data at the end of the first four sessions; therefore, our model contains four observations of WAI, but five observations of BDI, as WAI at session four predicts BDI at session five. In the RI-CLPM model, observed variables are parsed into within- and between-person components. This is illustrated in Figure 1; the top of the figure shows the observed BDI scores, each of which is predicted by the latent random intercept of BDI (the between-person component). Each observation of BDI is also predicted by its own latent within-person variable. Observed WAI scores are treated similarly. On the within-person level, we model the autoregressive effects of within-person BDI and within-person WAI, respectively. Additionally, the within-person variables are correlated at the first session, while their error terms are correlated at sessions two, three, and four. The primary effects of interest are the cross-lagged associations between within-person BDI and within-person WAI. Within-person BDI at a given session predicts within-person WAI at the end of the same session, while within-person WAI predicts within-person BDI at the start of the next session. We refer to these cross-lagged parameters as predicting *regressed change* in BDI and WAI because we controlled for the prior values of these variables (see Hamaker et al., 2015 and Cohen, Cohen, West, & Aiken, 2003). We did not assess change using difference scores.

The initial model was run with the following parameters constrained to be equal: the grand means of BDI and WAI, the factor loadings of the latent variables (which were constrained to 1), the autoregressive parameters for within-person WAI and for within-person BDI, the parameters predicting within-person BDI from within-person WAI, the parameters predicting within-person WAI from within-person BDI, and the covariances between the residuals of the within-person variables.

The estimated relation between two variables may be inflated if there is a linear change over time in both (Wang & Maxwell, 2015). Because we observed linear trends in BDI and WAI over time, we also evaluated a model that accounted for these trends (i.e., a detrended model). Detrended models may inflate the possibility of type II error in cases where removing variability related to the linear trends removes some of the variability in the relationship that the model is intended to estimate. Therefore, we present the results of both our primary model without detrending and a detrended model (see Falkenström, Solomonov, & Rubel, 2020). The detrended model followed the modeling approach described in Curran, Howard, Bainter, Lane, and McGinley (2014). This model, the latent curve model with

structured residuals, is equivalent to our RI-CLPM, but with the addition of a latent slope factor predicting WAI and a latent slope factor predicting BDI. The parameters that were constrained to be equal in the initial RI-CLPM were also constrained to be equal in the detrended model, with the exception of the grand means, which must be allowed to freely vary for the model to be detrended (for example code see the online supplement).

Results

Correlations among variables at each session are provided in Table 1. These are raw correlations, reflecting the combination of within- and between-patient variability. BDI and WAI were not significantly related to each other at sessions 1 and 2 but show a small, yet significant correlation at sessions 3 and 4. Means and SDs for BDI and WAI at each session are also reported in the table. We also calculated intraclass correlation coefficients to determine how much variability in scores was between-patient. The remaining variation is either within-patient or error. Most of the variance for both WAI total scores and BDI-II scores was between-person (ICCs of .79 and .75, respectively).

On average, patients reported an improvement in BDI-II scores of 5.37 points ($SD = 8.82$) between sessions 1 and 5 ($t(170) = 7.96, p < .001, d = -0.6$). Patients also reported an increase in WAI scores of 2.1 points ($SD = 7.76$) from session 1 to session 4 ($t(168) = -3.5, p < .001, d = 0.28$).

RI-CLPM: Modifications

The initial model did not fit well: $\chi^2(34, N = 191) = 128.23, p < .000$, root mean square error of approximation (RMSEA) = .12, 90% confidence interval (CI) [.10, .14], probability RMSEA .05 = .00, comparative fit index (CFI) = .94, standardized root mean square residual (SRMR) = .10, Akaike information criterion (AIC) = 10897.27. Therefore, as suggested by modification indices (MI; with the minimum MI value set to 5), all of the grand means of BDI were freed (MI value for BDI1 = 24.76; MI value for BDI5 = 6.37). This suggested modification is consistent with our finding that BDI scores changed over time. Freeing the means of the observed variables represents one of the changes required for detrending a model; it allows the model to separate out fluctuations in the group means from the cross-lagged parameters. This model accounts for changes in the overall means of the variables, while a detrended model adjusts for individual-specific trajectories (Falkenström et al., 2019). The fit of the resulting model was better: $\chi^2(30, N = 191) = 73.01, p < .000$, RMSEA = .09, 90% CI [.06, .11], probability RMSEA .05 = .01, CFI = .97, SRMR = .07, AIC = 10850.06. A chi-square difference test also suggested this model fit the data better: $\chi^2(4, N = 191) = 55.22, p < .001$. In this test, the null hypothesis is the simpler model (the model with more degrees of freedom). A low p -value suggests that this simpler model should be rejected in favor of the more complex model. Therefore, we rejected the initial model in favor of the model with the grand means of BDI estimated freely.

RI-CLPM: Results

Within-person WAI significantly predicted within-person BDI, with an increase of 1 point in within-person WAI predicting a 0.31 decrease in within-person BDI ($SE = 0.05, 95\%$

CI [-0.40, -0.22], $p < .001$). The standardized estimates ranged from -0.28 to -0.34; these estimates vary slightly because each standardized estimate is calculated using the standard deviation of the predictor and dependent variable, and this changes at each session. Regressed change in BDI also predicted regressed change in WAI, with a 1-point increase in within-person BDI predicting a 0.62 decrease in WAI (SE = 0.19, 95% CI [-0.30, -0.94], $p = .001$). Standardized estimates ranged from -.54 to -.64.¹

Detrending: Model Details and Results

The detrended model had acceptable fit: $\chi^2(20, N = 191) = 40.23, p = .005$, RMSEA = .07, 90% CI [.04, .11], probability RMSEA .05 = .12, CFI = .99, SRMR = .07, AIC = 10837.28. Regressed change in within-person WAI did not significantly predict regressed change in within-person BDI (estimate = -0.00, SE = 0.08, 95% CI [-0.14, 0.13], $p = .96$). Regressed change in BDI also did not predict regressed change in WAI (estimate = 0.38, SE = 0.43, 95% CI [-0.32, 1.08], $p = .37$).

Discussion

When we tested a reciprocal model of the alliance-outcome relationship without detrending, we found the alliance predicted subsequent symptom change and symptom change predicted subsequent alliance. The model we used allowed us not only to examine reciprocal relationships between our variables, but also to estimate these relations using within-person variability specifically. Further, our model allowed us to control for prior assessments of depressive symptoms. However, in a detrended model, we did not observe significant relationships between the alliance and symptom change. The alliance-outcome relationship we observed was modest and much of the alliance scores appeared attributable to patient factors. These findings are consistent with the view that the alliance has limited impact on outcomes on average in CBT for depression. The relationship we identified was not robust to detrending, suggesting further reason for caution.

At first glance, it may appear that similar estimates to ours are available in the meta-analytic literature. When one looks at all available studies of the alliance-outcome relationship without regard for their methodological features, the average relationship appears to be about .28 (Flückiger, Del Re, Wampold, & Horvath, 2018). However, model differences are substantial enough to raise questions about the extent to which these estimates can be meaningfully compared. Our estimate reflects the portion of the alliance that is not explained by stable between-person differences. Meta-analytic estimates have largely been based on raw scores, driven by a combination of within- and between-person variability, and thus, are fundamentally different than estimates that only utilize within-person variability.

¹Although we focused on the total WAI score, we also conducted exploratory analyses on the two WAI subscales (relate and agree) suggested by an exploratory factor analysis (Andrusyna, Tang, DeRubeis, & Luborsky, 2001). The results mirrored those of our primary analyses. For the relate subscale, BDI significantly predicted WAI-relate (median of standardized estimates = -0.44) and WAI-relate significantly predicted BDI (median of standardized estimates = -0.22). For the agree subscale, BDI significantly predicted WAI-agree (median of standardized estimates = approx. -0.60) and WAI-agree significantly predicted BDI (median of standardized estimates = -0.27).

This makes it especially important to consider how our RI-CLPM estimate of the alliance-outcome relationship compares to other findings using the same analytic approach. When estimating the relation of alliance and subsequent symptoms using the RI-CLPM, Falkenström et al. (2019) reported standardized estimates ranging from .16 to .21. Our estimates of a similar relationship ranged from .27 to .32, reflecting a slightly stronger alliance-outcome relationship. Our sample was composed exclusively of patients with depression and our outcome measure was focused on depressive symptoms specifically. In contrast, Falkenström and colleagues' (2019) sample was composed of a wider range of disorders and included a variety of psychotherapies. Despite these differences, the results of their detrended model mirrored ours: they found that alliance did not predict symptom change and symptom change did not predict alliance. This raises the question of whether the alliance and outcome may appear related only because both variables display linear change, and not because of a causal relationship between the two. An alternative possibility is that detrending our data was an overly conservative approach that may have resulted in type II error. The results of our detrended model, while not conclusive, cast doubt on the robustness of the within-person alliance-outcome relationship in the context of CBT of depression. When considering the magnitude of the effects we identified and the degree to which the alliance appeared attributable to patient factors, our findings leave us skeptical that variability in the alliance contributes substantially to the therapeutic benefits of CBT of depression.

The amount of within-person variability in the alliance was considerably smaller than the portion of between-patient variability (79% of variance was between-patient). The variability in the therapeutic alliance was largely accounted for by differences between patients, who had stable tendencies in their alliance scores. This finding is consistent with previous research finding that 67 – 69% of the variance in alliance scores on two subscales of the WAI was attributable to between patient differences in a trial of CBT of depression (Sasso, Strunk, Braun, DeRubeis & Brotman, 2016). Our estimates are somewhat higher than those of Falkenström, Granström, and Holmqvist (2013), who found that 54% of the variance observed in the patient-reported WAI was between-patient variability among patients participating in psychotherapy in a primary care setting in Sweden. Within- and between-patient variability in alliance scores is not frequently reported in the alliance literature, limiting our ability to assess how our estimates compare with other studies. Regardless, in the studies we identified which did report it, the majority of variability in alliance was between-patient. Barber and colleagues (2014) provided a particularly powerful demonstration of the patient's role in determining the alliance. They found patients' pre-treatment expectations for the alliance strongly predicted alliance ratings after patients met their therapists and began treatment (r s of around .50). Thus, at least in some samples, patient differences appear to account for a considerable portion of the variation in alliance scores.

Our findings suggest the alliance appears to be determined by patients to a large degree, the magnitude of the alliance-outcome relationship is only modest, and reciprocal modeling of the alliance and outcome only provides support for the alliance as a predictor of outcome when the model does not include detrending. Efforts to foster the alliance remain quite popular (see Muran & Barber, 2011). Although we are aware of no arguments against

fostering a positive alliance, our findings lead us to conclude that either other variables or a more complex mix of a number of factors are likely to be needed to provide a more satisfactory account of determinants of patients' therapeutic outcomes. However, it is important to note that these inferences are based on an analysis of the early sessions of a naturalistic study of CBT for depression. It may be that the alliance plays a more important role among some patients or when a more diverse set of intervention approaches are used (Hofmann & Hayes, 2019; Zilcha-Mano, 2018).

Limitations

There are several limitations worthy of note. First, although our model has a number of positive methodological features, even such complex models do not establish a causal relationship with the certainty of experimental designs. Second, the time between measurements varied, with the time from a symptom assessment to the next alliance assessment being about one hour (i.e., from the beginning to the end of a session), whereas the time from an alliance assessment to the next symptom assessment was several days (i.e., at the following session). Although we think our time lags are reasonable, it is often difficult to know which time lags are most appropriate for the effects of interest. Third, we modified our model to improve fit. Such modifications come with a risk of capitalizing on chance associations and generating sample specific findings. Although we tried to be thoughtful in our model modifications, we cannot rule out this possibility. Lastly, it is important for future research to explore whether relations of the alliance and outcome vary across different cultural contexts.

Conclusion

We found support for a reciprocal model of the alliance–outcome relationship in which alliance change predicts symptom change and symptom change predicts alliance change during early sessions of CBT of depression. Our findings were obtained in a model with a number of methodologically desirable features. Nonetheless, the alliance is to a large degree determined by patients, the magnitude of the alliance–outcome relationship is only modest, and support for the reciprocal alliance–outcome relationships was limited to the model that did not account for linear change in these variables. We encourage further research examining whether there are particular patients for whom the alliance may be more consequential.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This project was supported by the National Center for Research Resources (Award number TL1RR025753).

References

- Adler AD, Strunk DR, & Fazio RH. (2015). What changes in cognitive therapy for depression? An examination of cognitive therapy skills and maladaptive beliefs. *Behavior Therapy*, 46, 96–109. doi: 10.1016/j.beth.2014.09.001 [PubMed: 25526838]
- Allison PD. (1990). Change scores as dependent variables in regression analysis. *Sociological Methodology*, 93–114. doi: 10.2307/271083
- Allison PA. (2014). *Fixed effects regression methods for longitudinal data using SAS*. Cary, NC: SAS Institute.
- Allison P. (2015). Don't put lagged dependent variables in mixed models. Retrieved from <http://m.statisticalhorizons.com/?url=http%3A%2F%2Fstatisticalhorizons.com%2Fflagged-dependent-variables&width412>
- Andrusyna TP, Tang TZ, DeRubeis RJ, & Luborsky L. (2001). The factor structure of the working alliance inventory in cognitive-behavioral therapy. *The Journal of Psychotherapy Practice and Research*, 10, 173–178. [PubMed: 11402080]
- Barber JP, Zilcha-Mano S, Gallop R, Barrett M, McCarthy KS, & Dinger U. (2014). The associations among improvement and alliance expectations, alliance during treatment, and treatment outcome for major depressive disorder. *Psychotherapy Research*, 24, 257–268. doi: 10.1080/10503307.2013.871080 [PubMed: 24392793]
- Beck AT, Steer RA, & Brown BK. (1996). *Beck Depression Inventory manual* (2nd ed.). San Antonio, TX: Psychological Corporation.
- Cohen J, Cohen P, West SG, & Aiken LA. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3rd Edition. Mahwah, NJ: Routledge.
- Curran PJ, & Bauer DJ. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, 62, 583–619. doi: 10.1146/annurev.psych.093008.100356
- Curran PJ, Howard AL, Bainter SA, Lane ST, & McGinley JS. (2014). The separation of between-person and within-person components of individual change over time: A latent curve model with structured residuals. *Journal of Consulting and Clinical Psychology*, 82, 879–894. doi: 10.1037/a0035297 [PubMed: 24364798]
- Ebner-Priemer U, & Trull T. (2009). Ecological momentary assessment of mood disorders and mood dysregulation. *Psychological Assessment*, 21, 463–475. doi: 10.1037/a0017075 [PubMed: 19947781]
- Eubanks CF, Muran JC, & Safran JD. (2018). Alliance rupture repair: A meta-analysis. *Psychotherapy*, 55, 508–519. doi: 10.1037/pst0000185508 [PubMed: 30335462]
- Falkenström F, Finkel S, Sandell R, Rubel JA, & Holmqvist R. (2017). Dynamic models of individual change in psychotherapy process research. *Journal of Consulting and Clinical Psychology*, 85, 537–549. doi: 10.1037/ccp0000203 [PubMed: 28394170]
- Falkenström F, Granström F, & Holmqvist R. (2013). Therapeutic alliance predicts symptomatic improvement session by session. *Journal of Counseling Psychology*, 60, 317–328. doi:10.1037/a0032258317 [PubMed: 23506511]
- Falkenström F, Granström F, & Holmqvist R. (2014). Working alliance predicts psychotherapy outcome even while controlling for prior symptom improvement. *Psychotherapy Research*, 24, 146–159. doi: 10.1080/10503307.2013.847985 [PubMed: 24188079]
- Falkenström F, Kuria M, Othieno C, & Kumar M. (2019). Working alliance predicts symptomatic improvement in public hospital-delivered psychotherapy in Nairobi, Kenya. *Journal of Consulting and Clinical Psychology*, 87, 46–55. doi: 10.1037/ccp0000363 [PubMed: 30431300]
- Falkenström F, Solomonov N, & Rubel J. (2020). Using time-lagged panel data analysis to study mechanisms of change in psychotherapy research: Methodological recommendations. *Counselling and Psychotherapy Research*. Advance online publication. doi: 10.1002/capr.12293
- Feeley M, DeRubeis RJ, & Gelfand LA. (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. doi: 10.1037//0022-006x.67.4.578 [PubMed: 10450629]

- First MB, Spitzer RL, Gibbon M, & Williams JBW. (2002). Structured clinical interview for DSM-IV-TR axis I disorders, research version, patient edition. (SCID-I/P). New York: Biometrics Research, New York State Psychiatric Institute.
- Fisher H, Atzil-Slonim D, Bar-Kalifa E, Rafaeli E, & Peri T. (2016). Emotional experience and alliance contribute to therapeutic change in psychodynamic therapy. *Psychotherapy*, 53, 105–116. doi: 10.1037/pst0000041 [PubMed: 26785358]
- Flückiger C, Del Re AC, Wampold BE, & Horvath AO. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy*, 55, 316–340. doi: 10.1037/pst0000172 [PubMed: 29792475]
- Hamaker EL, Kuiper RM, & Grasman RPPP. (2015). A critique of the cross-lagged panel Model. *Psychological Methods*, 20, 102–116. doi: 10.1037/a0038889 [PubMed: 25822208]
- Hofmann SG, & Hayes SC. (2019). The future of intervention science: Process-based therapy. *Clinical Psychological Science*, 7, 37–50. doi: 10.1177/2167702618772296 [PubMed: 30713811]
- Horvath AO, & Greenberg LS. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, 36, 223–233. doi: 10.1037/0022-0167.36.2.223
- Lorenzo-Luaces L, DeRubeis RJ, & Webb CA. (2014). Client characteristics as moderators of the relation between the therapeutic alliance and outcome in cognitive therapy for depression. *Journal of Consulting and Clinical Psychology*, 82, 368–373. doi: 10.1037/a0035994 [PubMed: 24547921]
- Luborsky L. (1994). Therapeutic alliances as predictors of psychotherapy outcomes: Factors explaining the predictive success. In Horvath AO and Greenberg LS (Eds.) *The working alliance: Theory, research, and practice* (pp. 38–50). John Wiley & Sons: New York.
- Martin DJ, Garske JP, & Davis MK. (2000). Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, 68, 438–450. doi: 10.1037/0022-006X.68.3.438 [PubMed: 10883561]
- Muran JC, & Barber JP (Eds.). (2011). *The therapeutic alliance: An evidence-based guide to practice*. Guilford: New York.
- Muthén LK, & Muthén BO. (2017). *Mplus User's Guide*. Eighth Edition. Los Angeles, CA: Muthén & Muthén.
- Rubel JA, Rosenbaum D, & Lutz W. (2017). Patients' in-session experiences and symptom change: Session-to-session effects on a within-and between-patient level. *Behaviour Research and Therapy*, 90, 58–66. doi: 10.1016/j.brat.2016.12.007 [PubMed: 27998800]
- Sasso KE, Strunk DR, Braun JD, DeRubeis RJ, & Brotman MA. (2016). A re-examination of process-outcome relations in cognitive therapy for depression: Disaggregating within-patient and between-patient effects. *Psychotherapy Research*, 26, 387–398. doi: 10.1080/10503307.2015.1026423 [PubMed: 25876795]
- Schmidt ID, Pfeifer BJ, & Strunk DR. (2019). Putting the “cognitive” back in cognitive therapy: Sustained cognitive change as a mediator of in-session insights and depressive symptom improvement. *Journal of Consulting and Clinical Psychology*, 87, 446–456. doi: 10.1016/j.brat.2010.03.011 [PubMed: 30998047]
- Siev J, Huppert JD, & Chambless DL. (2009). The dodo bird, treatment technique, and disseminating empirically supported treatments. *the Behavior Therapist*, 32, 69–76.
- Tracey TJ, & Kokotovic AM. (1989). Factor structure of the Working Alliance Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 1, 207–210. doi: 10.1037/1040-3590.1.3.207
- Wang LP, & Maxwell SE. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods*, 20, 63–83. doi: 10.1037/met0000030 [PubMed: 25822206]
- Webb CA, Beard C, Auerbach RP, Menninger E, & Björgvinsson T. (2014). The therapeutic alliance in a naturalistic psychiatric setting: Temporal relations with depressive symptom change. *Behaviour Research and Therapy*, 61, 70–77. doi: 10.1016/j.brat.2014.07.015 [PubMed: 25156322]
- Xu H, & Tracey TJG. (2015). Reciprocal influence model of working alliance and therapeutic outcome over individual therapy course. *Journal of Counseling Psychology*, 62, 351–359. doi: 10.1037/cou0000089 [PubMed: 25985186]

- Zilcha-Mano S. (2018). Major developments in methods addressing for whom psychotherapy may work and why. *Psychotherapy Research*, 29, 693–708. doi: 10.1080/10503307.2018.1429691 [PubMed: 29409394]
- Zilcha-Mano S, Dinger U, McCarthy KS, & Barber JP. (2014). Does alliance predict symptoms throughout treatment, or is it the other way around? *Journal of Consulting and Clinical Psychology*, 82, 931–935. doi: 10.1037/a0035141 [PubMed: 24274627]
- Zilcha-Mano S, & Errázuriz P. (2015). One size does not fit all: Examining heterogeneity and identifying moderators of the alliance–outcome association. *Journal of Counseling Psychology*, 62, 579–591. doi: 10.1037/cou0000103 [PubMed: 26376176]
- Zilcha-Mano S, Eubanks CF, & Muran JC. (2019). Sudden gains in the alliance in cognitive behavioral therapy versus brief relational therapy. *Journal of Consulting and Clinical Psychology*, 87, 501–509. doi: 10.1037/ccp0000397 [PubMed: 31008637]

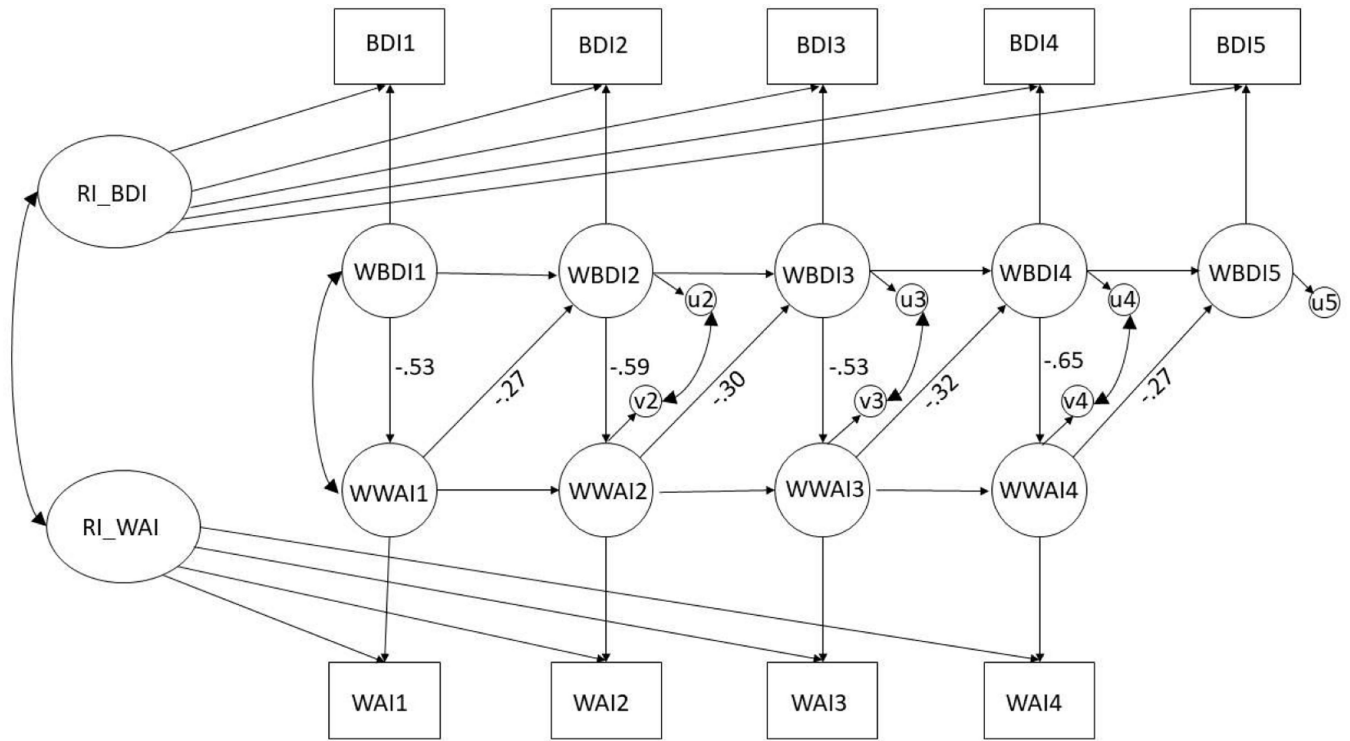


Figure 1. The Random Intercept Cross-Lagged Panel Model (RI-CLPM).

[We have included the original file linked as author guidelines request. However, these files were not displaying correctly in the assembled pdf file (despite multiple attempts). Therefore, we have also included the jpg file shown above to ensure that reviews will be able to see a correctly appearing version.]

Note. BDI stands for Beck Depression Inventory-II and WAI stands for Working Alliance Inventory. RI is an abbreviation for “random intercept”. W in front of a variable name denotes the within-person component of that variable. Error terms are represented by u and v. This diagram shows the decomposition of observed BDI and WAI scores into within- and between-person components in the RI-CLPM model and the standardized estimates of the effect of WBDI on WWAI and WWAI on WBDI at each time point.

Table 1.

Means, Standard Deviations, and Correlations among Beck Depression Inventory and Working Alliance Inventory Scores.

	BDI1	BDI2	BDI3	BDI4	BDI5	WAI1	WAI2	WAI3	WAI4
BDI2	.79*								
BDI3	.75*	.85*							
BDI4	.71*	.83*	.85*						
BDI5	.70*	.74*	.79*	.84*					
WAI1	.04	.00	-.03	-.00	-.04				
WAI2	.05	.02	-.07	-.03	-.07	.78*			
WAI3	-.02	-.03	-.16*	-.14	-.18*	.78*	.83*		
WAI4	.01	-.06	-.15	-.15*	-.19*	.73*	.81*	.88*	
Mean	27.9	26.15	24.39	23.04	22.60	58.17	58.57	59.20	60.12
SD	10.13	11.01	11.23	1.46	12.22	10.28	10.56	2.50	11.09

Note. Sample sizes range from 172 – 190. BDI = Beck Depression Inventory; WAI = Working Alliance Inventory.