

Journal of Counseling Psychology

Does the Working Alliance Mediate the Effect of Routine Outcome Monitoring (ROM) and Alliance Feedback on Psychotherapy Outcomes? A Secondary Analysis From a Randomized Clinical Trial

Heidi Brattland, John Morten Koksvik, Olav Burkeland, Christian A. Klöckner, Mariela Loreto Lara-Cabrera, Scott D. Miller, Bruce Wampold, Truls Ryum, and Valentina Cabral Iversen

Online First Publication, January 31, 2019. <http://dx.doi.org/10.1037/cou0000320>

CITATION

Brattland, H., Koksvik, J. M., Burkeland, O., Klöckner, C. A., Lara-Cabrera, M. L., Miller, S. D., Wampold, B., Ryum, T., & Iversen, V. C. (2019, January 31). Does the Working Alliance Mediate the Effect of Routine Outcome Monitoring (ROM) and Alliance Feedback on Psychotherapy Outcomes? A Secondary Analysis From a Randomized Clinical Trial. *Journal of Counseling Psychology*. Advance online publication. <http://dx.doi.org/10.1037/cou0000320>

Does the Working Alliance Mediate the Effect of Routine Outcome Monitoring (ROM) and Alliance Feedback on Psychotherapy Outcomes? A Secondary Analysis From a Randomized Clinical Trial

Heidi Brattland and John Morten Koksvik
Norwegian University of Science and Technology and St. Olavs
University Hospital, Trondheim, Norway

Olav Burkeland
St. Olavs University Hospital, Trondheim, Norway

Christian A. Klöckner
Norwegian University of Science and Technology

Mariela Loreto Lara-Cabrera
St. Olavs University Hospital, Trondheim, Norway

Scott D. Miller
The International Center for Clinical Excellence,
Chicago, Illinois

Bruce Wampold
Modum Bad Psychiatric Center, Vikersund, Norway, and
University of Wisconsin–Madison

Truls Ryum
Norwegian University of Science and Technology

Valentina Cabral Iversen
Norwegian University of Science and Technology and St. Olavs
University Hospital, Trondheim, Norway

Little is known about the mechanisms through which routine outcome monitoring (ROM) influences psychotherapy outcomes. In this secondary analysis of data from a randomized clinical trial (Brattland et al., 2018), we investigated whether the working alliance mediated the effect of the Partners for Change Outcome Monitoring System (PCOMS), a ROM system that provides session-by-session feedback on clients' well-being and the alliance. Adult individuals ($N = 170$) referred for hospital-based outpatient mental health treatment were randomized to individual psychotherapy either with the PCOMS ROM system, or without (treatment as usual [TAU]). Treatment was provided by the same therapists ($N = 20$) in both conditions. A multilevel mediation model was developed to test if there was a significant indirect effect of ROM on client impairment at posttreatment through the alliance at 2 months' treatment controlled for first-session alliance. Alliance ratings increased more from session 1 to 2 months' treatment in the ROM than TAU condition, and alliance increase was associated with less posttreatment impairment. A significant indirect effect of ROM on treatment outcomes through alliance increase ($p = .043$) explained an estimated 23.0% of the effect of ROM on outcomes. The results were consistent with a theory of the alliance as one mechanism through which ROM works.

Heidi Brattland and John Morten Koksvik, Department of Mental Health, Norwegian University of Science and Technology, and Department of Tiller DPS, St. Olavs University Hospital, Trondheim, Norway; Olav Burkeland, Department of Tiller DPS, St. Olavs University Hospital; Christian A. Klöckner, Department of Psychology, Norwegian University of Science and Technology; Mariela Loreto Lara-Cabrera, Department of Tiller DPS, St. Olavs University Hospital; Scott D. Miller, The International Center for Clinical Excellence, Chicago, Illinois; Bruce Wampold, Research Institute, Modum Bad Psychiatric Center, Vikersund, Norway, and Department of Counseling Psychology, University of Wisconsin–Madison; Truls Ryum, Department of Psychology, Norwegian University of Science and Technology; Valentina Cabral Iversen, Department of Mental Health, Norwegian University of Science and Technology, and Department of Tiller DPS, St. Olavs University Hospital.

John Morten Koksvik is now at Department of DPS Fredrikstad, Østfold Hospital, Fredrikstad, Norway.

Scott D. Miller holds the copyright to the Outcome and Session Rating Scales used in this study.

This work was supported by the Liaison Committee for Education, Research and Innovation in Central Norway under Grant 46056812 (653510 IPH). We are grateful to the Tiller Mental Health Center, St. Olavs University Hospital, and the Department of Research and Development, Norwegian University of Science and Technology for cofunding this research, and to all participating clients and participants. The authors report no conflicts of interest.

This research was presented at the 49th International Annual Meeting for the Society for Psychotherapy Research in Amsterdam, the Netherlands, June 27–30, 2018.

Correspondence concerning this article should be addressed to Heidi Brattland, Department of Mental Health, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Pb 8905, 7491 Trondheim, Norway. E-mail: heidi.brattland@ntnu.no

Public Significance Statement

Previous research indicates that psychotherapy clients may benefit more from treatment when their session-by-session levels of well-being and experiences of the alliance are monitored with short questionnaires. The results of this study suggest that some of this effect may be due to improvements in the collaborative working relationship over time.

Keywords: client feedback systems, common factors, mediation, routine outcome monitoring, working alliance

Routine outcome monitoring (ROM) is a family of client self-report questionnaires developed to track psychotherapy clients' session-by-session levels of well-being, symptoms, or functioning to help therapists detect problems in a client's response to treatment as it evolves (Howard, Moras, Brill, Martinovich, & Lutz, 1996; Lambert, 2007).¹ Several meta-analyses have found ROM to improve treatment outcomes (Hansen & Lambert, 2003; Knaup, Koesters, Schoefer, Becker, & Puschner, 2009; Lambert & Shimokawa, 2011; Østergård, Randa, & Hougaard, 2018; Shimokawa, Lambert, & Smart, 2010). However, a more complex picture of ROM's influence on treatment outcomes is beginning to emerge. A recent Cochrane review (Kendrick et al., 2016) deemed the available evidence for ROM insufficient. The heterogeneity in results between studies is substantial (Østergård et al., 2018). Some studies have reported null-findings (e.g., Davidsen et al., 2017; Hansson, Rundberg, Österling, Öjhagen, & Berglund, 2013; Rise, Eriksen, Grimstad, & Steinsbekk, 2016) and others, indications of adverse effects of ROM (de Jong, Segaar, Ingenhoven, van Busschbach, & Timman, 2018; Errázuriz & Zilcha-Mano, 2018; van Oenen et al., 2016). The mixed findings highlight the need for a more nuanced understanding of how the process works, a question which, to date, has received limited research attention (Wampold, 2015). Identifying mechanisms of change inherent in ROM could help improve the clinical implementation of these interventions and thus, maximize their effect. The present study investigated one potential change mechanism for ROM, the working alliance, in a secondary analysis of data from randomized clinical trial.

The contextual feedback intervention theory (Sapyta, Riemer, & Bickman, 2005) posits that providing therapists with negative feedback (i.e., information about a discrepancy between their current performance and some desired goal, such as helping the client improve) motivates corrective action. This is thought to be especially true if the feedback is direct, specific to the therapists' behaviors, promptly delivered, and comes from a credible source. ROM systems are designed to fulfill these requirements and provide information that therapists would otherwise have difficulties obtaining (Lambert & Shimokawa, 2011). As such, ROM systems are theorized to work through correcting cognitive biases that prevent therapists from detecting treatment failures (Macdonald & Mellor-Clark, 2015). Beyond prompting corrective action in therapists, however, this theory does not specify nor explain through what processes ROM works to avoid negative outcomes. One possibility is that ROM mobilizes the common factors of psychotherapy and particularly, the working alliance (Miller, Hubble, Duncan, & Wampold, 2010), whose association to treatment outcomes is well documented (Horvath & Bedi, 2002; Horvath, Del

Re, Flückiger, & Symonds, 2011; Horvath & Symonds, 1991; Martin, Garske, & Davis, 2000).

There is reason to believe that ROM might influence the quality of the alliance. Paying close attention to whether therapy is helpful for the client, and taking immediate action when this is not the case, is in itself likely to facilitate and even enhance a collaborative working relationship. Moreover, some ROM systems provide feedback to therapists on the alliance as well as on client impairment and well-being. For instance, the Partners for Change Outcome Management System (PCOMS; Miller, Duncan, Sorrell, & Brown, 2005), which was used in the present trial, contains an alliance measure that therapists administer and discuss with their clients toward the end of each treatment session. Thus, the PCOMS is designed to identify alliance problems and facilitate the continuous work of negotiating the nature of the collaborative working relationship (Bordin, 1979; Hatcher & Barends, 2006; Safran & Muran, 2006). As such, we would expect its use to improve the alliance over the course of treatment. Possibly, this could explain some of the positive effect of the PCOMS on treatment outcomes that was found in several studies (e.g., Anker, Duncan, & Sparks, 2009; Reese, Norsworthy, & Rowlands, 2009; Reese, Toland, Slone, & Norsworthy, 2010; Schuman, Slone, Reese, & Duncan, 2015; She et al., 2018; Slone, Reese, Mathews-Duvall, & Kodet, 2015; see Østergård et al., 2018, for a thorough review and discussion of the PCOMS literature).

One way to explore the alliance as a change mechanism for ROM is to investigate if it mediates ROM's effect. To date, no ROM study has tested a mediation hypothesis. There is however evidence that (1) alliances improve more over the course of treatment for clients receiving treatment with ROM and alliance feedback than for those receiving treatment without ROM (Janse, De Jong, Van Dijk, Hutschemaekers, & Verbraak, 2017; McClintock, Perlman, McCarrick, Anderson, & Himawan, 2017), and (2) alliance improvements from one session to the next predict lower impairment levels in subsequent sessions (e.g., Crits-Christoph, Gibbons, Hamilton, Ring-Kurtz, & Gallop, 2011; Falkenström, Ekelbad, & Holmqvist, 2016; Zilcha-Mano, Dinger, McCarthy, & Barber, 2014; Zilcha-Mano & Errázuriz, 2015; Zilcha-Mano et al., 2016). As discussed by Zilcha-Mano (2017), the temporal precedence of alliance improvement to symptom change found in the majority of studies to investigate this issue, contradicts claims that the alliance is merely a byproduct of symptom reduction (e.g., Barber, 2009; DeRubeis, Brotman, & Gibbons, 2005). Nevertheless, it is not clear in what capacity the alliance is involved in

¹ Other terms are client feedback systems, patient-reported outcome measures (PROMS), and feedback-informed treatment (FIT).

therapeutic change processes. The alliance, or the broader therapeutic relationship, could be curative in itself; for instance, the process of negotiating the alliance and repairing ruptures could provide corrective emotional experiences, insight, and interpersonal learning (e.g., Norcross, 2002; Safran & Muran, 2000). Alternatively, the alliance could be a precondition for other mechanisms to cause client change (e.g., Hatcher & Barends, 2006). For instance, a stronger agreement on tasks and goals could allow therapists choose more suitable interventions, or those interventions could have a larger impact in the context of a stronger emotional bond. The alliance could work as a mechanism of change for ROM in both these capacities.

The present study builds on a previous publication from a randomized clinical trial in which a positive effect of the PCOMS ROM system on treatment outcomes was established (Brattland et al., 2018). Correcting for therapist variability and initial distress in a multilevel regression model (clients nested within therapists), a small ($d = 0.26$), but significant ($p = .037$) superior effect on treatment outcome was found for ROM over TAU. Presently, we hypothesized that some of the difference between ROM and TAU in the clients' posttreatment impairment would be explained by higher alliance ratings in the ROM than TAU condition at 2 months' treatment, controlling for first-session alliance levels (i.e., that the alliance would mediate the effect of ROM on treatment outcomes).

Method

Participants

Clients. The trial took place in a general psychiatric outpatient department at a Norwegian hospital-based mental health clinic, which serves a population of adult clients (age 18 years or older) with moderate to severe mental health problems of all diagnostic categories.

Participants were recruited from the waitlist for individual outpatient treatment at the clinic. The exclusion criterion was inability to complete questionnaires (e.g., because of illiteracy, very low cognitive functioning, or poor understanding of the Norwegian language); all other individuals accepted for individual outpatient treatment were eligible for participation in the trial.

The final sample (excluding nine no-show clients, see the following text) consisted of 161 clients. The majority ($n = 100$, 63.3%) were female, and their mean age was 34.1 years old. About half of the clients ($n = 74$, 46.5%) were single, but most clients were living with someone ($n = 135$, 84.9%) and reported having someone in whom they could confide ($n = 129$, 82.2%). Half of the sample were in active employment ($n = 77$, 50%) and the majority had education beyond primary school ($n = 131$, 82.9%).

Client diagnoses were made by their therapists, at the outset of therapy, who used the M.I.N.I. International Neuropsychiatric Interview (Sheehan et al., 1998) and International Statistical Classification of Diseases and Health Related Problems (ICD-10; World Health Organization, 2010). Diagnoses were for clinical purposes only and the reliability of the diagnostic process was not assessed. The most common diagnostic categories in this sample were affective ($n = 59$, 30.1%) and anxiety disorders ($n = 59$, 30.1%), followed by hyperkinetic disorders (e.g., attention-deficit/hyperactivity disorder; $n = 20$, 10.2%), personality disorders ($n =$

17, 8.7%), and other diagnoses ($n = 19$, 9.7%), including psychotic and eating disorders. Thirty-five (17.9%) were diagnosed with two disorders, and 22 (11.2%) were diagnosed with none. No information regarding characteristics of the population was available to assess the representativeness of the sample.

Therapists. Treatment was provided by 20 therapists (16 women and four men; 11 clinical psychologists, six psychiatrists, and three other mental health care professionals). Staff experience with PCOMS ROM system ranged from 1 month to 5 years. On a seven-point Likert scale (1 = *very little*, 7 = *very much*), therapists reported being most influenced by psychodynamic therapy models ($Mdn = 6$; range = 2–7), followed by humanistic/existential ($Mdn = 5$; range = 1–6) and cognitive ($Mdn = 4$; range = 2–7) models. Therapists provided treatment for clients in both the ROM and TAU conditions.

Conditions

Treatment as usual (TAU). In the control condition, clients received nonmanualized outpatient individual psychotherapy. Therapists were responsible for assessing and diagnosing their clients, as well as determining the treatment approach, frequency of sessions, and treatment length. All cases were discussed in interdisciplinary teams. TAU clients attended a mean of 13.01 sessions ($SD = 10.92$, $Mdn = 10$, range = 1–54). No information about treatment content was gathered. However, because the same therapists treated clients in both conditions, any differences in outcomes between conditions would likely be due to the experimental intervention rather than systematic differences in the content of the treatment.

Routine outcomes monitoring (ROM). Participants in the experimental condition received the same type of outpatient individual psychotherapy as those in the TAU condition, with the addition of the PCOMS measures every session. ROM clients in this sample attended a mean of 12.04 sessions ($SD = 9.35$, $Mdn = 9$, range = 1–45).

As discussed in Brattland et al. (2018), the clinic underwent an extensive implementation process over the course of this trial. Therapists received regular training and supervision in the use of the PCOMS according to the International Center for Clinical Excellence's manuals on feedback-informed treatment (FIT; Bertolino & Miller, 2012). The PCOMS measures were administered on computer tablets, using a web-based program called *FIT outcomes* (www.fit-outcomes.com). In brief, the procedure is as follows: Therapists administer a self-report measure of well-being, the Outcome Rating Scale (ORS; Miller, Duncan, Brown, Sparks, & Claud, 2003), to their clients during the first few minutes of every treatment session. The ORS consists of four items (symptoms, relational functioning, social role functioning, and global well-being), each resulting in a score from 0 (*minimal well-being*) to 10 (*maximum well-being*). The total score is then plotted on a graph. This graph contains scores from previous session as well as expected trajectories of change, which are based on normative data for clients with the same or similar initial ORS scores (Miller, 2011). Consequently, lack of improvement is immediately detectable to both the client and the therapist. When scores fall below the expected treatment response, therapists are instructed to discuss this with the client and explore ways in which the treatment approach can be adjusted.

Of particular relevance to the present study, during the final minutes of each visit, clients complete a measure of the therapeutic alliance, the Session Rating Scale (SRS; Duncan et al., 2003). Like the ORS, the sum of SRS's four items (therapeutic relationship, goals and topics, approach or method, and overall experience of the alliance) is displayed on the graph. Therapists are instructed not to aim for "perfect" alliance scores, but rather to use the SRS as a tool to facilitate honest feedback about the therapeutic process and to respond in a nondefensive, cooperative manner. Only with clients in the ROM condition were the ORS and the SRS administered on a session-by-session basis, and these scores were not included in the data material for the present study.

Fidelity. The protocol dictated that therapists administer the PCOMS measures to clients only in the ROM condition and never to clients randomized to TAU. At each client's treatment termination, their therapists were asked, in a self-report questionnaire, if they had in fact administered the ORS and the SRS to that client. We received 60 responses for clients in the TAU condition; here, the PCOMS measures had been never administered to 59 clients and every session to one client. In the ROM condition, we received 58 responses and the measured were reportedly administered every session to 51 clients, some sessions to two clients, and never to five clients. These data indicate that the PCOMS measures were administered or withheld according to the protocol for all cases but six. The reason for this nonadherence to the protocol for six cases is not clear.

Measures

Impairment. The Behavior and Symptoms Identification Scale (BASIS-32; Eisen, Wilcox, Leff, Schaefer, & Culhane, 1999) measured symptoms and psychosocial functioning at pre- and posttreatment. BASIS-32 consists of 32 items that are rated on a five-point Likert scale (0 = *no difficulty*; 4 = *extreme difficulty*), generating five subscale scores (relation to self/others, daily living/role functioning, depression/anxiety, impulsive/addictive behavior, and psychosis) and an overall mean score. The latter was used in this analysis. BASIS-32 has been validated in several studies (Doerfler, Addis, & Moran, 2002; Eisen et al., 1999; Hoffmann, Capelli, & Mastrianni, 1997; Jerrell, 2005; Klinkenberg, Cho, & Vieweg, 1998; Russo et al., 1997); it has been found sensitive to change and moderately correlated to other measures of symptoms and function. In the present study, the internal consistency was high as indicated by Cronbach's alphas of .934 at pretreatment and .958 at posttreatment.

Alliance. The quality of the working alliance was assessed at Session 1 (T1) and at 2 months' treatment (T2) with the short version of the Working Alliance Inventory (WAI-S; Horvath & Greenberg, 1989; Tracey & Kokotovic, 1989), a 12-item questionnaire based on Bordin (1979)'s three working alliance dimensions: Emotional bond and agreement on the goals and tasks of therapy. Items are scored on a seven-point Likert scale, ranging from 1 (*strongly disagree*) to 7 (*strongly agree*), with higher scores indicating better working alliance. WAI-S is widely used in research. Despite a well-established reliability in previous studies (e.g., Busseri & Tyler, 2003; Hanson, Curry, & Bandalos, 2002; Horvath, 1994; Tracey & Kokotovic, 1989), the item-total correlation for two items with reversed wording (specifically, two of the four items that load on the goal dimension of the working alliance) was

low in the present study, .259 and .082, respectively. Following the recommendations of Field (2013), these items were removed, with a resulting improvement in internal consistency (i.e., Cronbach's alpha increased from .899 to .936 at T1 and from .949 to .955 at T2).

Recruitment and Procedure

The trial was conducted within the daily practice at a hospital mental health outpatient clinic. An intake team, which consisted of health personnel who were not part of the research team, assessed all treatment referrals both for suitability for treatment and for eligibility to participate in the trial. This assessment was based on referral letters, which typically contained a brief description of the presenting problems and relevant medical or psychiatric history.

Prospective participants were assigned to therapists before inclusion to the trial. They then met with one of the principal investigators to give informed consent, complete baseline measures, and be randomized into the ROM or TAU condition. The randomization was performed using a web-based randomization program for medical research (<https://webcrf.medisin.ntnu.no>) and a 1:1 allocation ratio. It was not practically feasible to blind participants, therapists or investigators to the results of the randomization.

The outcomes measures were pen-and-paper questionnaires. Data from baseline, T1, T2, and treatment termination were used; the first two were completed at the clinic and the remaining questionnaires were mailed to participants. Questionnaires were re-sent twice if participants failed to return them. The procedure was approved by the Regional Committee for Research Ethics (Case number 2011/1711). The trial was registered on Clinical Trials (clinicaltrials.gov; identifier: NCT01796223).

We had originally planned to recruit 120 participants but increased the sample size to 170, due to concerns that missing data (see the following text) would decrease the statistical power. Recruitment took place over 3 years, during which time a total of 1,655 clients were referred for treatment (see Figure 1). More individuals are referred to the clinic than it has the capacity to serve and consequently, a substantial number of referrals (typically, those who are considered to be suitable for other treatment format such as group therapy or lower level counseling services) are not offered individual outpatient treatment. Additionally, during periods of the study, the intake team forgot to assess referrals for eligibility to the trial. A total of 659 individuals were invited to participate via mail and telephone. The remaining 489 individuals did not respond to the invitation, declined participation, or started treatment before inclusion. Nine participants did not attend any sessions and were discharged without treatment, leaving 161 clients in the final sample.

As in many other naturalistic studies, there was missing data in this trial. In Figure 1, the number of cases with missing data at each point of measure is depicted. Note that the total missing rather than cumulative data is presented, in contrast to the flowchart in Brattland et al. (2018). As such, there is some overlap in the missing columns as some clients had missing data on more than one measure. In total, 70 clients (43.5%) failed to return one or more measure. The mediation model estimated several different relationships between variables (see the Data Analysis section) and cases were included if they had sufficient data to estimate at least one of these relationships, as is typical for multilevel models

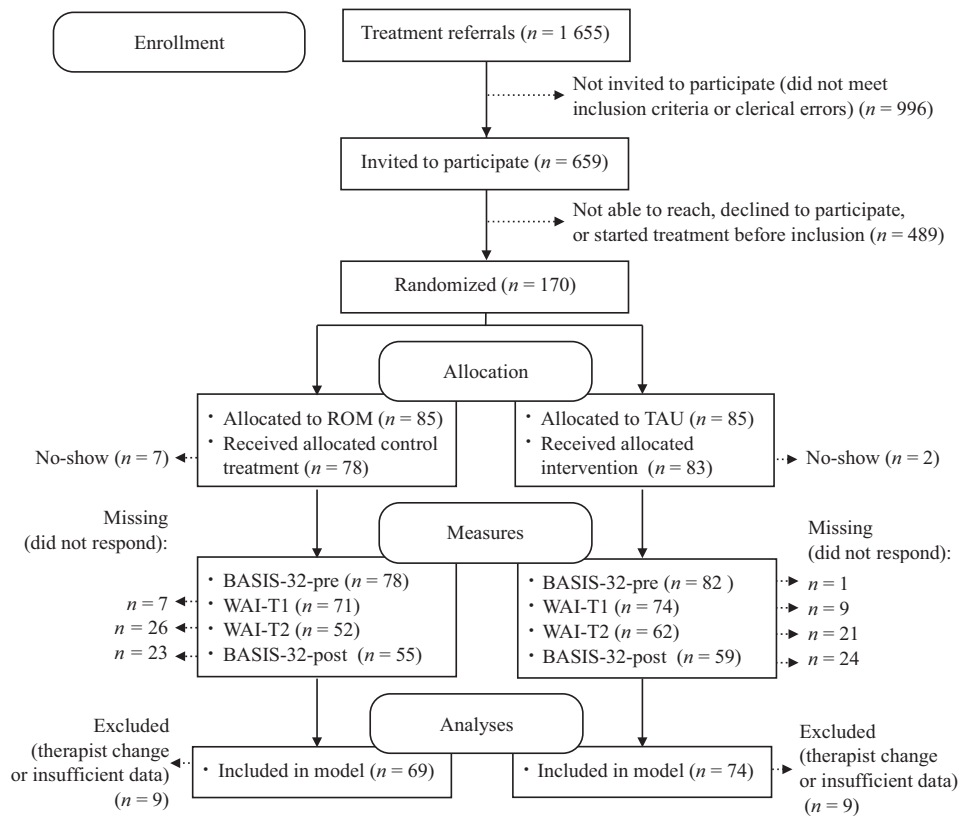


Figure 1. Participant flowchart. TAU = treatment as usual condition; ROM = routine outcome monitoring condition; T1 = first session; T2 = 2 months' treatment. Total missing data per measurement time is displayed; these overlap somewhat (i.e., the same client may have missing data at more than one measure).

(MLMs). No clients had missing data at all four measures, but eight (5.0%) had missing data on three measures (T1, T2, and posttreatment) and consequently, were not included in the mediation model. In addition, 10 clients had missing data at the therapist level due to a change of therapist mid-treatment. These were also excluded from the model. Ultimately, in the mediation model, 143 cases were included.

Data Analysis

General analytic strategy. Controlling for the nested structure or shared covariance between clients treated by the same therapist (Adelson & Owen, 2012; Wampold & Serlin, 2000), our hypotheses were tested in an MLM (Snijders & Bosker, 2012) with clients at Level 1 nested within therapists at Level 2. This was done using the statistical software Mplus (Muthén & Muthén, 1998–2017). In small samples, a nonparametric distribution is expected for the indirect effects in mediation models. Because of this, we used Bayesian estimation, which makes no assumptions about the prior distribution but instead, uses a posterior distribution based on the observed values (Hamaker & Klugkist, 2011; Robert, 2007). The Bayesian analyses were performed with 30,000 iterations. We did not provide any prior values to the model. Bayesian posterior trace plots for each parameter were inspected to determine if the models converged. We report median point estimates and posterior

standard deviations for the parameters as well as credibility intervals (CrIs), which are based on the percentile points of the posterior probability distribution and describe the range in which the true values of the parameters are likely to fall. The R^2 is reported to indicate how much variance each model explained, the intraclass correlation coefficient (ICC) to indicate the proportion of this variance that was due to differences between therapists, and the deviance information criteria (DIC) to indicate model fit.

Test of hypotheses. A random intercepts, fixed slopes, two-level mediation model was set up as follows: The X variable was condition (coded 0 for TAU and 1 for ROM), the mediator M , the alliance (WAI-T2 controlled for WAI-T1), and the dependent variable Y , treatment outcome (BASIS-32-post controlled for BASIS-32-pre). All variables were measured at the client level. In one step and within the same model, we tested the effect of ROM on the alliance (by convention referred to as path a), the effect of alliance on treatment outcome (path b), the indirect effect of ROM on outcome via the alliance (the product of paths a and b), and the residual direct effect of ROM on treatment outcome when the indirect effect was controlled for (path c').

The planned mediation model was expressed at Level 1 by the following two equations:

$$Y_{ij} = b_{0j} + b_1 \text{BASIS-32-pre}_{ij} + b_2 \text{WAI-T2}_{ij} + c' \text{Condition}_{ij}$$

$$M_{ij} = a_{0j} + a_1 \text{WAI-T1} + a_2 \text{Condition}_{ij}$$

where Y_{ij} is the posttreatment BASIS-32 score for client i treated by therapist j ; b_{0j} is the intercept for therapist j , b_1 is the slope estimate for BASIS-32-pre; b_2 for WAI-T2; c' for condition; M_{ij} is the WAI-T2 score for client i treated by therapist j ; a_{0j} is the intercept for therapist j , a_1 is the slope estimate for the covariate WAI-T1; and a_2 for condition. Here, the indirect effect is expressed in $a_2 b_2$ and the residual direct effect, in c' . At the therapist level, the intercepts b_0 and a_0 for therapist j is expressed by the equations

$$b_{0j} = \gamma_{Y00} + \mu_{Y0j} \quad \text{and} \quad a_{0j} = \gamma_{M00} + \mu_{M0j}$$

where γ_{Y00} is the mean intercept for Y ; μ_{Y0j} is the Y intercept residual for therapist j ; γ_{M00} is the mean intercept for M ; and μ_{M0j} is the M intercept residual for therapists j .

In line with contemporary mediation frameworks (e.g., Hayes, 2009; Hayes & Rockwood, 2017), mediation was supported if the estimate for the indirect effect $a_2 b_2$ was statistically different from zero. The total effect c was inferred from the sum of the indirect and direct effects (i.e., $c = a_2 b_2 + c'$), as recommended by Kenny, Korchmaros, and Bolger (2003) and others. To aid the interpretation of the indirect effect, we report the partially standardized indirect effect size (Preacher & Kelley, 2011), which is defined as the ratio of the indirect effect ab to the standard deviation of the outcome variable (i.e., $a_2 b_2 / SD_{\text{BASIS-32-post}}$). The partially standardized indirect effect size is interpreted as the number of standard deviations by which the outcome variable is expected to increase or decrease per unit change in the mediator of size a . We also report the somewhat dated, but more intuitively interpretable, extent of mediation effect size, which is given by $a_2 b_2 / c$.

The a priori hypothesis in this trial was directional: A negative indirect effect of ROM on posttreatment impairment through alliance increase. Accordingly, we report one-sided significance tests with an alpha level of .05. We also report 90% CrIs, of which the upper or lower bound (depending on whether the effect is hypothesized to be positive or negative) represent the value below or above which we would expect 95% of future observations to fall. As only one side of the CrIs are of interest when hypotheses are directional, the resulting error rate is 5% (Pocock, 2003).

Handling of missing data. We mitigated the effect of missing data as follows. First, as described in the preceding text, MLMs

included all cases with sufficient data to estimate at least one parameter. Thus, only 18 (11.2%) cases were excluded from the mediation model, and data from all remaining cases informed the estimation of one or several parameters in the model. Second, we compared cases with missing data points to those with complete data to assess whether there were systematic differences between these groups that needed to be accounted for in the final model. Third, although the main results are based on observed values only, we also tested the mediation hypothesis on a data set in which missing data had been imputed. For this, we used the maximum likelihood imputation procedure (Schafer & Graham, 2002), with all other observed variables as auxiliary variables. The resulting estimates are presented in the text.

Results

Preliminary Analyses

Conditions were similar on all baseline variables (all $ps > .05$) except for social network; significantly more clients in the ROM than TAU condition reported having nobody in whom they could confide, $\chi^2(1) = 4.826, p = .028$. Social network was found to have no influence neither on posttreatment impairment nor on the alliance and consequently, we did not include this variable in the final mediation model.

The mean scores on all variables are presented in Table 1. Unexpectedly, ROM clients had significantly lower WAI-T1 scores, which we had planned to model as a covariate to WAI-T2 only. Because WAI-T1 was measured after the first session (i.e., the first "dose" of treatment), this could reflect a negative effect of ROM on early alliance. Consequently, we adjusted the planned mediation model by modeling condition as a covariate to WAI-T1 (referred to in Table 2 as parameter d_1) so as to control for this relationship in the final model. We also performed a post hoc analysis of the influence of early alliance on treatment outcome (see the following text).

Clients with missing data points were similar to those with complete data sets in the distribution of diagnoses and the baseline variables age, gender, and social network, but had higher pretreatment impairment scores (mean ROM = 1.673, $SD = 0.743$; mean TAU = 1.379, $SD = 0.588$; $t(158) = -2.789, p = .06$), more frequently reported having no higher education than primary

Table 1
The Mean Behavior and Symptoms Identification Scale (BASIS-32) and Working Alliance Inventory (WAI) Scores in Total and Per Condition

Variable	<i>n</i>	Total <i>M (SD)</i>	Condition		<i>p</i>
			TAU <i>M (SD)</i>	ROM <i>M (SD)</i>	
BASIS-32-pre	160	1.51 (.67)	1.43 (.63)	1.58 (.71)	.181
BASIS-32-post	114	.91 (.68)	.98 (.69)	.84 (.66)	.280
WAI-T1	145	5.00 (1.26)	5.22 (1.25)	4.77 (1.25)	.033
WAI-T2	114	5.06 (1.34)	4.91 (1.39)	5.24 (1.27)	.192

Note. p values are obtained through independent samples t tests. BASIS-32-post = Behavior and Symptoms Identification Scale scores at posttreatment; BASIS-32-pre = baseline Behavior and Symptoms Identification Scale scores; condition = treatment as usual (TAU; coded 0) or routine outcome monitoring (ROM; coded 1); WAI-T1 = first-session Working Alliance Inventory scores; WAI-T2 = Working Alliance Inventory scores at 2 months' treatment.

Table 2
Two-Level Mediation Model of the Effects of Routine Outcome Monitoring (ROM) on Posttreatment Behavior and Symptoms Identification Scale (BASIS-32) Scores Mediated by Working Alliance Inventory (WAI) Scores at 2 Months' Treatment

Dependent variable	Parameter	Estimate (SD)	CrI	p
Fixed effects				
WAI-T2	Intercept a_0	2.148 (.473)	1.393, 2.943	<.001
	WAI-T1 a_1	.528 (.084)	.387, .661	<.001
	Condition a_2	.490 (.212)	.143, .840	.011
WAI-T1	BASIS-32-pre a_3	-.377 (.182)	-.677, -.081	.019
	Intercept d_0	5.240 (.153)	4.987, 5.493	<.001
BASIS-32-post	Condition d_1	-.505 (.219)	-.867, -.148	.011
	Intercept b_0	1.332 (.187)	1.038, 1.651	<.001
	BASIS-32-pre b_1	.466 (.102)	.299, .632	<.001
	WAI-T2 b_2	-.098 (.052)	-.183, -.011	.032
	Condition c'	-.140 (.115)	-.329, .045	.080
	Indirect effect a_2b_2	-.043 (.034)	-.112, -.001	.043
Variances				
WAI-T2	Within therapists	1.059 (.171)	.826, 1.381	<.001
	Between therapists	.224 (.221)	.040, .686	<.001
	ICC	.175		
BASIS-32-post	Within therapists	.320 (.050)	.252, .415	<.001
	Between therapists	.066 (.061)	.018, .195	<.001
	ICC	.171		

Note. All estimates are unstandardized regression coefficients. BASIS-32-post = Behavior and Symptoms Identification Scale scores at posttreatment; BASIS-32-pre = baseline Behavior and Symptoms Identification Scale scores (grand mean centered); CrI = 90% credibility intervals; condition = treatment as usual (TAU; coded 0) or routine outcome monitoring (ROM; coded 1); ICC = intraclass correlation coefficient; SD = Posterior standard deviation; WAI-T1 = first-session Working Alliance Inventory scores; WAI-T2 = Working Alliance Inventory scores at 2 months' treatment.

school (63.0% v. 37.0%; $\chi^2(1) = 5.274, p = .022$), and were more often single (56.8% v. 43.2%; $\chi^2(1) = 10.059, p = .002$), and living alone (66.7% v. 33.3%; $\chi^2(1) = 6.231, p = .013$). This indicates that clients with missing data points had lower levels of functioning and consequently, that data was not missing completely at random. Of these variables, only pretreatment impairment was found to significantly predict posttreatment impairment and the alliance. Because of this, we made a second adjustment to the planned model by adding a parameter for the effect of pretreatment impairment on WAI-T2 (referred to in Table 2 as parameter a_3) as well as on BASIS-32-post (b_1 , as originally planned). We also investigated the potential confounding effects of demographic variables in a post hoc analysis (see the following text).

Meditation Analysis

The results of the mediation analysis are presented in Table 1 and Figure 2. Controlled for WAI-T1 and BASIS-32-pre, ROM was associated with higher WAI-T2 ratings (parameter a_2). Controlled for BASIS-32-pre and condition, higher WAI-T2 ratings were associated with lower BASIS-32-post scores (parameter b_2). The indirect effect of ROM through the alliance (a_2b_2) went in the same direction as the residual direct effect (c'), and was significantly different from zero. We calculated the total effect c of ROM on posttreatment impairment ($a_2b_2 + c'$) to be -0.183 . The partially standardized effect size in this model (based on $SD_{BASIS-32-post} = 0.686$) was -0.063 and the extent of mediation, 0.235. This model explained 39.5% of the variance in

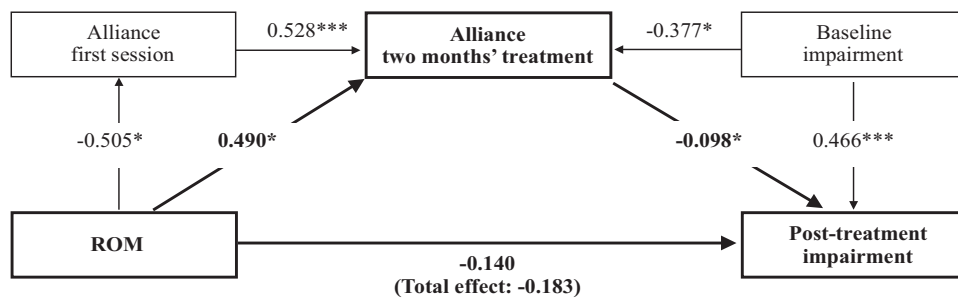


Figure 2. Mediation model. ROM = routine outcome monitoring condition; Alliance = Working Alliance Inventory (WAI) scores; impairment = Behavior and Symptoms Identification Scale (BASIS-32) scores. Bold lines and letters represent the parameters that are relevant for the mediation hypothesis. * $p < .05$. *** $p < .001$.

BASIS-32-post and 34.2% of the variance in WAI-T2 as indicated by the R^2 . As indicated by the ICC, differences between therapists explained 17.1% of the variance in BASIS-32-post and 17.5% of the variance in WAI-T2. The DIC was 963.776.

With imputation of missing values, all parameters went in the same direction as those obtained in the observed values only data set, and a_2b_2 , a_2 , and b_2 were all statistically significant. The estimate a_2b_2 for the indirect effect of ROM on posttreatment impairment through the mediator was -0.024 ($SD = 0.018$, 90% CrI $[-0.061, -0.002]$, $p = .032$), the estimate a_2 for the effect of ROM on the mediator, 0.327 ($SD = 0.145$, 90% CrI $[0.088, 0.566]$, $p = .012$), the estimate b_2 for the effect of the mediator on posttreatment impairment, -0.082 ($SD = 0.040$, 90% CrI $[-0.148, -0.016]$, $p = .021$), and the residual direct effect c' of ROM on posttreatment impairment, -0.091 ($SD = 0.081$, 90% CrI $[-0.226, 0.039]$, $p = .127$). This model explained 34.1% of the variance in BASIS-32-post and 39.9% of the variance in WAI-T2 as indicated by the R^2 . As indicated by the ICC, 17.0% of the variance in BASIS-32-post and 17.5% of the variance in WAI-T2 was explained by differences between therapists.

Post Hoc Analysis

Effect of first-session alliance on treatment outcomes. ROM had a negative effect on first-session alliance. Due to this unexpected finding we investigated if WAI-T1 influenced treatment outcome by modeling this variable as a predictor to BASIS-32-post while retaining all parameters from the mediation model. This was not the case; the estimate for the effect of first-session alliance on posttreatment impairment was negligible and nonsignificant, -0.015 ($SD = 0.066$, 90% CrI $[-0.123, 0.093]$, $p = .413$), and model fit was slightly worse (DIC = 965.523) than that of the previous mediation model. Thus, the negative effect of ROM on first-session alliance did not influence treatment outcomes.

Therapist variability. The therapist ICCs for posttreatment distress and alliance were high (see Table 2), indicating substantial differences between therapists in both the mediator and the outcome. To explore this issue further we tested series of random intercept, random slope models where all parameters from the mediation model were retained and the slope estimates for parameters a_2 , b_2 , and c' were allowed to vary at random, one by one, between therapists. There was a significant between-therapist variability in both the effect of ROM on the alliance (mean $a_2 = 0.474$; variance = 0.308, $SD = 0.381$, 90% CrI $[0.033, 1.135]$, $p > .001$), the effect of the alliance on posttreatment impairment (mean $b_2 = -0.078$; variance = 0.002, $SD = 0.002$, 90% CrI $[0.919, 1.508]$, $p < .001$), and the residual direct effect of ROM on posttreatment impairment (mean $c' = -0.139$; variance = 0.048, $SD = 0.076$, 90% CrI $[0.004, 0.213]$, $p < .001$).

Potential biases. The final sample included data from 10 clients that attended treatment less than 2 months (i.e., the point in time at which WAI-T2 was measured). This is potentially problematic; a stronger statement can be made about mediation if a timeline can be established in which a mediator is measured before the outcome. As these cases had missing data at WAI-T2, they were not included in the estimation of parameters a_2 and b_2 and consequently, retaining them in the model did not bias the estimation of the indirect effect a_2b_2 . However, six of these clients had returned the posttreatment questionnaires and consequently, were included in the estimation of parameter c' , which could bias the direct and total effect as well as extent of

mediation. To investigate if this were the case we tested the mediation model on a subsample of cases that excluded these 10 clients. The resulting direct effect was $c' = -0.138$ ($SD = 0.114$, 90% CrI $[-0.329, 0.049]$, $p = .110$), very similar to the corresponding estimate in the complete sample (see Table 2), which suggests that including data from clients who had quit treatment prior to the measurement of the mediator did not bias our results.

Also included in the final sample was data from six cases for which therapists had not adhered to protocol, either by administering the PCOMS measures to clients in the TAU condition, or not administering those measures to clients in the ROM condition. As is common in RCTs, these six cases were included in the main analysis; applying a post hoc selection criterion to the sample would reduce the external validity. To investigate whether these six cases biased the results, we tested the mediation model on a subsample in which the cases were omitted. This resulted in a slightly higher estimate for the indirect effect ($a_2b_2 = -0.047$, $SD = 0.040$, 90% CrI $[-0.126, 0.000]$, $p = .052$). Note that only 95 cases had sufficient data to estimate path ab here, which could explain the somewhat higher p value obtained for the indirect effect in this subsample. Regarding the individual paths in the mediation model, the effect of ROM on the alliance was stronger in this subsample ($a_2 = 0.545$, $SD = 0.216$, 90% CrI $[0.195, 0.905]$, $p = .007$), the effect of the alliance on treatment outcome was the same ($b_2 = -0.097$, $SD = 0.057$, 90% CrI $[-0.190, -0.002]$, $p = .045$), and the residual direct effect of ROM on treatment outcome, very similar ($c' = -0.144$, $SD = 0.119$, 90% CrI $[-0.338, 0.052]$, $p = .113$). Thus, retaining the cases in which therapists reported not having adhered to protocol did not appear to influence the results substantially.

To rule out the potentially confounding effects of third variables, we tested if any of the demographic variables (age, gender, marital status, level of education, work status, living situation, and social network) were correlated with the mediator or outcome variable. None of the demographic variables predicted WAI-T2 or BASIS-32-post (all $ps < .05$). Moreover, adding these variables as covariates to the mediation model did not substantially alter any of the parameters presented in Table 2. Thus, there was no evidence of confounding by any of the demographic variables that were measured in this trial.

Discussion

This secondary analysis of data from a naturalistic randomized clinical trial explored the working alliance as a mediator for the effect of the PCOMS, a ROM system which targets both treatment progress (i.e., session-by-session well-being) and the alliance. Clients' alliance ratings increased more in the ROM than TAU condition from T1 to T2, and more alliance increase predicted less posttreatment impairment. As hypothesized, there was a significant indirect effect of ROM through alliance increase on posttreatment impairment. Thus, the effect of ROM on treatment outcome was mediated by the alliance. As indicated by the partially standardized effect size, a ROM client experiencing the average alliance increase for that condition was predicted to score 0.063 standard deviations lower on the posttreatment impairment measure than a client with the average alliance increase for the TAU condition. An estimated 23.5% of the effect of ROM on outcomes was explained by the alliance increase that ROM clients experienced, as indicated by the extent of mediation effect size.

The data in this trial was consistent with a theory of the alliance as a mechanism through which some of the effects of ROM are transmitted (Miller et al., 2010). Two causal mechanisms are implicit in the theory: First, that therapists' use of ROM and alliance feedback causes the alliance to improve, and second, that the alliance improvement causes better treatment outcomes. Of course, causality is difficult to establish in psychotherapy research and demonstration of statistical mediation alone, and particularly from a single study, cannot support such claims (Kazdin, 2009). It is vital to consider our results in the context of theory and previous research.

Regarding the effect of ROM on the alliance, the contextual feedback intervention theory (Sapyta et al., 2005) predicts that regularly administered alliance measures facilitate the detection of alliance problems which may otherwise be difficult for therapists to recognize. Provided that therapists respond to this feedback in ways that help repair the alliance problems, the alliance should then improve over time. The feedback that the PCOMS's alliance measure provides matches well with the characteristics of feedback most likely to instigate therapist behavior change in this theory: It is specific to therapists' alliance-related behavior, delivered shortly after that behavior (i.e., at the end of every treatment session), and designed to give information that is relevant to improving the alliance.

The theory that ROM and alliance feedback may work to gradually strengthen the alliance over time is supported by the findings from previous research. Qualitative studies indicate that therapists actively use client feedback in general (Brattland et al., 2016) and ROM feedback in particular (e.g., Ones, Anderssen, Borg, & Karlsson, 2015; Snyder & Aafjes-van Doorn, 2016; Sundet, 2012) in their work to improve the collaborative relationship in therapy. Two controlled trials (Janse et al., 2017; McClintock et al., 2017) investigated the session-by-session development of the alliance for clients receiving treatment with and without ROM systems that included alliance feedback (the PCOMS and the Common Factors Feedback System, respectively). Similar to our results, ROM clients experienced more alliance increase than those receiving TAU in both these studies. One controlled trial (Rise, Eriksen, Grimstad, & Steinsbekk, 2012) found no effect of the PCOMS on alliance scores at six weeks but did not assess early alliance and consequently, was not able to investigate if the alliance increased over time. Thus, consistent with theory and previous research, our findings indicate that systematically tracking clients' treatment responses, including their experiences of the alliance, may facilitate the development of the working alliance in therapy.

Importantly however, the unique contribution of specific feedback on the alliance (as opposed to feedback on impairment or well-being) to treatment outcome could not be determined in the present trial. Previous findings have produced mixed findings regarding the benefits of alliance feedback. In a meta-analytic review, Lambert and Shiokawa (2011) found the effect of the Outcome Questionnaire-45 (OQ-45; Lambert, 2004) to be higher in studies in which the OQ-45 had been used in conjunction with Clinical Support Tools (CSTs), which include alliance feedback. However, none of the six studies included in this analysis specifically compared the use OQ-45 alone to the use of both OQ-45 and CSTs. More recent dismantling studies have not found alliance feedback to augment ROM's effect on treatment outcomes

(Mikeal, Gillaspay, Scoles, & Murphy, 2016; Errázuriz & Zilcha-Mano, 2018) or the working alliance (Reese et al., 2013). This suggests that merely providing alliance feedback to therapists may not be sufficient to improve the alliance. More than likely, what matters is how therapists respond to this information, a question which was not addressed in our trial.

The notion that alliance and relationship work can cause client change (e.g., Safran & Muran, 2000; Norcross, 2002) has, as discussed in the introduction, been the subject of some controversy. Our results do shed light on this issue. More alliance growth was associated better treatment outcomes, but because both variables were measured at only two points in time each, we were not able to establish a clear time line in which the alliance improved prior to and independent of changes in impairment levels. Accordingly, we cannot rule out the alternative explanation that the alliance improvement was a byproduct of, rather than a cause for, ROM clients' increasing well-being. Related to this, in the two previous studies that reported alliance growth with ROM and alliance feedback (Janse et al., 2017; McClintock et al., 2017), ROM did not improve the overall treatment outcomes. These findings are consistent with the notion that the alliance does not influence outcomes. If this is the case, then some other, unknown mechanisms than alliance improvement were responsible for the superior effects of ROM in the present trial (see Brattland et al., 2018). However, a temporal precedence of alliance improvement to symptom relief has been documented in several methodologically sophisticated studies (e.g., Crits-Christoph et al., 2011; Falkenström et al., 2016; Zilcha-Mano et al., 2014; Zilcha-Mano & Errázuriz, 2015; Zilcha-Mano et al., 2016), and neither Janse et al. (2017) nor McClintock et al. (2017) reported whether there was an association between the alliance and treatment outcomes in their studies. Consequently, it is possible that the alliance increase with ROM in these two studies in fact worked to improve treatment outcomes, but was counteracted by other mechanisms that had the opposite effect. Alternatively, the alliance increase might have had a stronger impact on outcomes in our study than in those by Janse et al. (2017) and McClintock et al. (2017), due to some unknown factor; accumulating evidence suggests that the magnitude of the alliance-outcome association may vary according to characteristics of clients, therapists, and treatments (e.g., Falkenström, Granström, & Holmqvist, 2013; Hoffart, Øktedalen, Langkaas, & Wampold, 2013; Zilcha-Mano & Errázuriz, 2015; Zilcha-Mano, Lipsitz, & Errázuriz, 2018; Zilcha-Mano, Muran, Hungr, Eubanks, Safran, & Winston, 2016).

A related debate is whether alliance work is therapeutic in itself (e.g., Safran & Muran, 2000; Norcross, 2002), or if the alliance is a precondition for other processes to cause client change (e.g., Hatcher & Barends, 2006). Our results are consistent with both views. That is, it is possible that clients in the ROM condition directly benefitted from the alliance improvements they experienced but also, for instance, that therapists chose more suitable intervention or techniques because of a progressively stronger agreement on goals or tasks (i.e., that therapist actions mediated the effect of the alliance on treatment outcome), or that stronger alliances increased the impact of these interventions on treatment outcome (i.e., that the alliance moderated the impact of therapist actions on outcome). This trial was not designed to address this issue. It is, in general, challenging to disentangle the alliance from actions in therapy; because the common factors are conceptualized

at a higher level of abstraction than therapist actions, the alliance does not exist independent of therapeutic interventions (Hatcher & Barends, 2006; Wampold & Imel, 2015).

An unexpected finding in this study was that ROM clients had lower first-session alliance scores than TAU clients. Speculatively, having been invited to reflect on any negative aspects of the alliance with a responsive therapist in the first session, ROM clients had a more considered and realistic view of the early alliance. This may have facilitated the communication about the treatment process from the first session onward and consequently, the negotiation of the goals and tasks of therapy (Bordin, 1979) as well as the repair of alliance ruptures (Safran & Muran, 2000). However, the observed differences in alliance scores between conditions may have other explanations as well. For instance, it may reflect a negative reaction to the PCOMS measures on the part of the clients, so that PCOMS initially impaired the formation of a good working alliance. If so, the alliance growth in the ROM condition may have reflected a “catching up” to the alliance level that TAU clients were at from the beginning of treatment. That said, as indicated in the post hoc analysis, the lower first-session alliance for ROM clients did not have a negative impact on treatment outcome. Similar findings were reported by Anker, Owen, Duncan, and Sparks (2010) and Owen, Miller, Seidel, and Chow (2016).

The therapist ICCs for both posttreatment distress and alliance were unusually high in this study. On both variables, about 17% of the variance was due to difference between therapists, compared with the 7% to 8%, which is typically reported in psychotherapy studies (Baldwin & Imel, 2013). Exploring this issue further in a post hoc analysis, we found that therapists significantly differed in the influence of ROM on the alliance, in the influence of the alliance on outcomes, and the residual direct effect of ROM on posttreatment. Of course, post hoc findings should be interpreted with care and in this instance, statistical power at the therapist level was poor. Nevertheless, in light of previous demonstrations of therapist differences in ROM effects (e.g., Anker et al., 2009; de Jong, van Sluis, Nugter, Heiser, & Spinhoven, 2012; Simon, Lambert, Harris, Busath, & Vazquez, 2012), this finding hints at different ways of working with ROM for different therapists.

Limitations

As discussed previously, a major limitation in this study was the inability to determine whether the observed alliance improvements in the ROM conditions temporally preceded and uniquely influenced clients' impairment levels. Nor were we able to disaggregate the within- and between-client components of the alliance. Consequently, this study does not add to the debate regarding the role of the alliance as a change mechanism in therapy. As argued by Hayes and Rockwood (2017), it is rare for studies in the social sciences to be able to study cause-effect relationships under optimal conditions, and these authors recommended that the various assumptions for causal inference be regarded as ideals or recommendations rather than literal requirements. Moreover, a single mediation analysis cannot alone prove or disprove a theory of causality, but previous research can help fill in the gaps left by methodological shortcomings (e.g., Kazdin, 2009). Our results are consistent with the theory that some of ROM's effects may be transmitted through alliance improvements, and in our reading of

the literature we find partial support for this theory. However, because of the unresolved issues regarding the specific nature of the influence on the alliance on outcomes, this interpretation is tentative.

Similar to other ROM studies, neither therapists nor clients were blinded to condition. Although the use of the same therapists in both conditions minimizes the impact of differences between therapists on outcomes, it increases the risk of the results being influenced by therapists' outcome expectations (e.g., performance bias). Also, as outcomes were assessed by client self-report, the results could have been influenced by client's knowledge of having been randomized to the experimental or TAU condition (assessment bias). This issue is difficult to bypass in ROM studies as therapists, and in most cases also clients, by necessity will know whether or not ROM is administered. Researcher allegiance could also bias our findings, as one of PCOMS' developers was a coauthor of this article.

Regarding external validity, like many other studies conducted in practice-setting there was a high proportion of missing data in this trial. Clients with missing data points had higher baseline impairment levels than those with complete data sets, indicating that data was not missing completely at random. We dealt with this by using all available data from all cases, controlling for impairment levels on both the mediator and the outcome, and testing the mediation model on a data set where missing values had been imputed. Imputation did not alter the conclusions drawn in this study. This suggests that if missing data biased the findings, then this was not related to any of the variables that were measured in this study. Nevertheless, the high proportion of missing data implies that some caution must be exercised in generalizing from our results. Also, a larger client sample would have increased the statistical power and consequently, our confidence in the results. Finally, the removal two items from the WAI-S, while reducing the risk of biases related to arbitrary scoring, also decreases the internal validity of our findings.

Implications

These limitations notwithstanding, this study adds to the scarce literature regarding the clinically important question of how ROM works. It is the first mediation analysis of ROM published to date and one of few to systematically investigate the influence of ROM on a purposed mechanism of change. The randomized controlled design increases our confidence in the purposed causal link between the use of ROM and alliance feedback, and the alliance. Another strength of this study was the use of independent process and outcome measures (i.e., not the PCOMS' alliance and well-being measures) to assess the effect of ROM; the use of ROM questionnaires to measure the impact of those same questionnaires, which has been the procedure in the majority of previous ROM studies, may compromise both the internal and external validity of findings.

Future studies might investigate other theoretically derived mediators as well as moderators at both the client and therapist level. Ideally studies should be designed so as to allow both an investigation of within- and between-client components in the relevant variables, and of the relationship between process and symptoms as it unfolds over time. We would also suggest that future research focus on how therapists work with these tools.

Ultimately, a more refined understanding of how ROM works could help improve the clinical use of these interventions so that more clients benefit from their use. Our results suggest that using alliance feedback to repair ruptures and improve the therapeutic relationship might increase the likelihood that clients will benefit from therapy. If replicated, this finding would support the priority that some ROM systems give to alliance feedback.

References

- Adelson, J. L., & Owen, J. (2012). Bringing the psychotherapist back: Basic concepts for reading articles examining therapist effects using multilevel modeling. *Psychotherapy: Theory, Research, & Practice, 49*, 152–162. <http://dx.doi.org/10.1037/a0023990>
- Anker, M. G., Duncan, B. L., & Sparks, J. A. (2009). Using client feedback to improve couple therapy outcomes: A randomized clinical trial in a naturalistic setting. *Journal of Consulting and Clinical Psychology, 77*, 693–704. <http://dx.doi.org/10.1037/a0016062>
- Anker, M. G., Owen, J., Duncan, B. L., & Sparks, J. A. (2010). The alliance in couple therapy: Partner influence, early change, and alliance patterns in a naturalistic sample. *Journal of Consulting and Clinical Psychology, 78*, 635–645. <http://dx.doi.org/10.1037/a0020051>
- Baldwin, S. A., & Imel, Z. E. (2013). Therapist effects: Findings and methods. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 256–297). Hoboken, NJ: Wiley.
- Barber, J. P. (2009). Toward a working through of some core conflicts in psychotherapy research. *Psychotherapy Research, 19*, 1–12. <http://dx.doi.org/10.1080/10503300802609680>
- 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. <http://dx.doi.org/10.1037/0022-006X.68.6.1027>
- Bertolino, B., & Miller, S. D. (2012). *ICCE manuals on feedback-informed treatment* (Vol. 1–6). Chicago, IL: ICCE Press.
- Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research, & Practice, 16*, 252–260. <http://dx.doi.org/10.1037/h0085885>
- Brattland, H., Høiseth, J. R., Burkeland, O., Inderhaug, T. S., Binder, P. E., & Iversen, V. C. (2018). Learning from clients: A qualitative investigation of psychotherapists' reactions to negative verbal feedback. *Psychotherapy Research, 28*, 545–559. <http://dx.doi.org/10.1080/10503307.2016.1246768>
- Brattland, H., Koksvik, J. M., Burkeland, O., Gråwe, R. W., Klöckner, C., Linaker, O. M., . . . Iversen, V. C. (2018). The effects of routine outcome monitoring (ROM) on therapy outcomes in the course of an implementation process. A randomized clinical trial. *Journal of Counseling Psychology*. Advance online publication. <http://dx.doi.org/10.1037/cou0000286>
- Busseri, M. A., & Tyler, J. D. (2003). Interchangeability of the working alliance inventory and working alliance inventory, short form. *Psychological Assessment, 15*, 193–197. <http://dx.doi.org/10.1037/1040-3590.15.2.193>
- 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*, 267–278. <http://dx.doi.org/10.1037/a0023668>
- Davidson, A. H., Poulsen, S., Lindschou, J., Winkel, P., Tróndarson, M. F., Waadegaard, M., & Lau, M. (2017). Feedback in group psychotherapy for eating disorders: A randomized clinical trial. *Journal of Consulting and Clinical Psychology, 85*, 484–494. <http://dx.doi.org/10.1037/ccp0000173>
- de Jong, K., Segaar, J., Ingenhoven, T., van Busschbach, J., & Timman, R. (2018). Adverse effects of outcome monitoring feedback in patients with personality disorders: A randomized controlled trial in day treatment and inpatient settings. *Journal of Personality Disorders, 32*, 393–413. http://dx.doi.org/10.1521/pepi_2017_31_297
- de Jong, K., van Sluis, P., Nugter, M. A., Heiser, W. J., & Spinhoven, P. (2012). Understanding the differential impact of outcome monitoring: Therapist variables that moderate feedback effects in a randomized clinical trial. *Psychotherapy Research, 22*, 464–474. <http://dx.doi.org/10.1080/10503307.2012.673023>
- DeRubeis, R. J., Brotman, M. A., & Gibbons, C. J. (2005). A conceptual and methodological analysis of the nonspecifics argument. *Clinical Psychology: Science and Practice, 12*, 174–183. <http://dx.doi.org/10.1093/clipsy.bpi022>
- Doerfler, L. A., Addis, M. E., & Moran, P. W. (2002). Evaluating mental health outcomes in an inpatient setting: Convergent and divergent validity of the OQ-45 and BASIS-32. *The Journal of Behavioral Health Services & Research, 29*, 394–403. <http://dx.doi.org/10.1007/BF02287346>
- Duncan, B. L., Miller, S. D., Sparks, J. A., Claud, D. A., Reynolds, L. R., Brown, J., & Johnson, L. D. (2003). The Session Rating Scale: Preliminary psychometric properties of a “working” alliance measure. *Journal of Brief Therapy, 3*, 3–12.
- Eisen, S. V., Wilcox, M., Leff, H. S., Schaefer, E., & Culhane, M. A. (1999). Assessing behavioral health outcomes in outpatient programs: Reliability and validity of the BASIS-32. *The Journal of Behavioral Health Services & Research, 26*, 5–17. <http://dx.doi.org/10.1007/BF02287790>
- Errázuriz, P., & Zilcha-Mano, S. (2018). In psychotherapy with severe patients discouraging news may be worse than no news: The impact of providing feedback to therapists on psychotherapy outcome, session attendance, and the alliance. *Journal of Consulting and Clinical Psychology, 86*, 125–139. <http://dx.doi.org/10.1037/ccp0000277>
- Falkenström, F., Ekeblad, A., & Holmqvist, R. (2016). Improvement of the working alliance in one treatment session predicts improvement of depressive symptoms by the next session. *Journal of Consulting and Clinical Psychology, 84*, 738–751. <http://dx.doi.org/10.1037/ccp0000119>
- Falkenström, F., Granström, F., & Holmqvist, R. (2013). Therapeutic alliance predicts symptomatic improvement session by session. *Journal of Counseling Psychology, 60*, 317–328. <http://dx.doi.org/10.1037/a0032258>
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). London, UK: SAGE.
- Hamaker, E. L., & Klugkist, I. (2011). Bayesian estimation of multilevel models. In J. Hox & J. K. Roberst (Eds.), *Handbook of advanced multilevel analysis*. Hove, UK: Psychology Press.
- Hansen, N. B., & Lambert, M. J. (2003). An evaluation of the dose-response relationship in naturalistic treatment settings using survival analysis. *Mental Health Services Research, 5*, 1–12. <http://dx.doi.org/10.1023/A:1021751307358>
- Hanson, W. E., Curry, K. T., & Bandalos, D. L. (2002). Reliability generalization of working alliance inventory scale scores. *Educational and Psychological Measurement, 62*, 659–673. <http://dx.doi.org/10.1177/0013164402062004008>
- Hansson, H., Rundberg, J., Österling, A., Öjehagen, A., & Berglund, M. (2013). Intervention with feedback using Outcome Questionnaire 45 (OQ-45) in a Swedish psychiatric outpatient population. A randomized controlled trial. *Nordic Journal of Psychiatry, 67*, 274–281. <http://dx.doi.org/10.3109/08039488.2012.736534>
- Hatcher, R. L., & Barends, A. W. (2006). How a return to theory could help alliance research. *Psychotherapy: Theory, Research, & Practice, 43*, 292–299. <http://dx.doi.org/10.1037/0033-3204.43.3.292>

- Hatcher, R. L., & Gillaspay, J. A. (2006). Development and validation of a revised short version of the Working Alliance Inventory. *Psychotherapy Research, 16*, 12–25. <http://dx.doi.org/10.1080/10503300500352500>
- Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs, 76*, 408–420. <http://dx.doi.org/10.1080/03637750903310360>
- Hayes, A. F., & Rockwood, N. J. (2017). Regression-based statistical mediation and moderation analysis in clinical research: Observations, recommendations, and implementation. *Behaviour Research and Therapy, 98*, 39–57. <http://dx.doi.org/10.1016/j.brat.2016.11.001>
- Hoffart, A., Øktedalen, T., Langkaas, T. F., & Wampold, B. E. (2013). Alliance and outcome in varying imagery procedures for PTSD: A study of within-person processes. *Journal of Counseling Psychology, 60*, 471–482. <http://dx.doi.org/10.1037/a0033604>
- Hoffmann, F. L., Capelli, K., & Mastrianni, X. (1997). Measuring treatment outcome for adults and adolescents: Reliability and validity of BASIS-32. *Journal of Mental Health Administration, 24*, 316–331. <http://dx.doi.org/10.1007/BF02832665>
- Horvath, A. O. (1994). *Empirical validation of Bordin's pantheoretical model of the alliance: The Working Alliance Inventory perspective*. New York, NY: Wiley.
- Horvath, A. O., & Bedi, R. P. (2002). The alliance. In C. Norcross (Ed.), *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients* (pp. 37–69). London, UK: Oxford University Press.
- Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy: Theory, Research, & Practice, 48*, 9–16. <http://dx.doi.org/10.1037/a0022186>
- Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology, 36*, 223–233. <http://dx.doi.org/10.1037/0022-0167.36.2.223>
- 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. <http://dx.doi.org/10.1037/0022-0167.38.2.139>
- Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy. Efficacy, effectiveness, and patient progress. *American Psychologist, 51*, 1059–1064. <http://dx.doi.org/10.1037/0003-066X.51.10.1059>
- Janse, P. D., De Jong, K., Van Dijk, M. K., Hutschemaekers, G. J. M., & Verbraak, M. J. P. M. (2017). Improving the efficiency of cognitive-behavioural therapy by using formal client feedback. *Psychotherapy Research, 27*, 525–538. <http://dx.doi.org/10.1080/10503307.2016.1152408>
- Jerrell, J. M. (2005). Behavior and symptom identification scale 32: Sensitivity to change over time. *The Journal of Behavioral Health Services & Research, 32*, 341–346. <http://dx.doi.org/10.1007/BF02291833>
- Kazdin, A. E. (2009). Understanding how and why psychotherapy leads to change. *Psychotherapy Research, 19*(4–5), 418–428. <http://dx.doi.org/10.1080/10503300802448899>
- Kendrick, T., El-Gohary, M., Stuart, B., Gilbody, S., Churchill, R., Aiken, L., . . . Moore, M. (2016). Routine use of patient reported outcome measures (PROMs) for improving treatment of common mental health disorders in adults. *Cochrane Database of Systematic Reviews, 7*(7), CD011119.
- Kenny, D. A., Korchmaros, J. D., & Bolger, N. (2003). Lower level mediation in multilevel models. *Psychological Methods, 8*, 115–128. <http://dx.doi.org/10.1037/1082-989X.8.2.115>
- Klinkenberg, W. D., Cho, D. W., & Vieweg, B. (1998). Reliability and validity of the interview and self-report versions of the BASIS-32. *Psychiatric Services, 49*, 1229–1231. <http://dx.doi.org/10.1176/ps.49.9.1229>
- Knaup, C., Koesters, M., Schoefer, D., Becker, T., & Puschner, B. (2009). Effect of feedback of treatment outcome in specialist mental healthcare: Meta-analysis. *The British Journal of Psychiatry, 195*, 15–22. <http://dx.doi.org/10.1192/bjp.bp.108.053967>
- Lambert, M. J. (2004). *Administration and scoring manual for the OQ-45.2 (Outcome Questionnaire)*. Wilmington, DE: American Professional Credentialing Services.
- Lambert, M. J. (2007). Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy Research, 17*, 1–14. <http://dx.doi.org/10.1080/10503300601032506>
- Lambert, M. J., Bailey, R., Kimball, K., Shimokawa, K., Harmon, S. C., & Slade, K. (2007). *Clinical support tools manual-brief version-40*. Salt Lake City, UT: OQ Measures.
- Lambert, M. J., & Shimokawa, K. (2011). Collecting client feedback. *Psychotherapy: Theory, Research, & Practice, 48*, 72–79. <http://dx.doi.org/10.1037/a0022238>
- Macdonald, J., & Mellor-Clark, J. (2015). Correcting psychotherapists' blindsidedness: Formal feedback as a means of overcoming the natural limitations of therapists. *Clinical Psychology & Psychotherapy, 22*, 249–257. <http://dx.doi.org/10.1002/cpp.1887>
- Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 68*, 438–450. <http://dx.doi.org/10.1037/0022-006X.68.3.438>
- McClintock, A. S., Perlman, M. R., McCarrick, S. M., Anderson, T., & Himawan, L. (2017). Enhancing psychotherapy process with common factors feedback: A randomized, clinical trial. *Journal of Counseling Psychology, 64*, 247–260. <http://dx.doi.org/10.1037/cou0000188>
- Mikeal, C. W., Gillaspay, J. A., Jr., Scoles, M. T., & Murphy, J. J. (2016). A dismantling study of the Partners for Change Outcome Management System. *Journal of Counseling Psychology, 63*, 704–709. <http://dx.doi.org/10.1037/cou0000168>
- Miller, S. D. (2011, November 22). *Cutting edge feedback*. Retrieved from <https://www.scottdmiller.com/cutting-edge-feedback/>
- Miller, S. D., Duncan, B. L., Brown, J., Sparks, J., & Claud, D. (2003). The Outcome Rating Scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy, 2*, 91–100.
- Miller, S. D., Duncan, B. L., Sorrell, R., & Brown, G. S. (2005). The partners for change outcome management system. *Journal of Clinical Psychology, 61*, 199–208. <http://dx.doi.org/10.1002/jclp.20111>
- Miller, S. D., Hubble, M. A., Chow, D., & Seidel, J. (2015). Beyond measures and monitoring: Realizing the potential of feedback-informed treatment. *Psychotherapy: Theory, Research, & Practice, 52*, 449–457. <http://dx.doi.org/10.1037/pst0000031>
- Miller, S. D., Hubble, M. A., Duncan, B. L., & Wampold, B. E. (2010). Delivering what works. In B. L. Duncan, S. D. Miller, B. E. Wampold, & M. A. Hubble (Eds.), *The heart and soul of change: Delivering what works in therapy* (2nd ed.). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/12075-014>
- Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Author.
- Norcross, J. C. (2002). *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients* (pp. 37–69). London, UK: Oxford University Press.
- Oanes, C. J., Anderssen, N., Borg, M., & Karlsson, B. (2015). How do therapists respond to client feedback? A critical review of the research literature. *Scandinavian Psychologist, 2*, e17. <http://dx.doi.org/10.15714/scandpsychol.2.e17>
- Østergård, O. K., Randa, H., & Hougaard, E. (2018). The effect of using the Partners for Change Outcome Management System as feedback tool in psychotherapy—A systematic review and meta-analysis. *Psychotherapy Research*. Advance online publication. <http://dx.doi.org/10.1080/10503307.2018.1517949>
- Owen, J., Miller, S. D., Seidel, J., & Chow, D. (2016). The working alliance in treatment of military adolescents. *Journal of Consulting and*

- Clinical Psychology*, 84, 200–210. <http://dx.doi.org/10.1037/ccp0000035>
- Pocock, S. J. (2003). The pros and cons of noninferiority trials. *Fundamental & Clinical Pharmacology*, 17, 483–490.
- Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods*, 16, 93–115. <http://dx.doi.org/10.1037/a0022658>
- Reese, R. J., Gillaspay, J. A., Jr., Owen, J. J., Flora, K. L., Cunningham, L. C., Archie, D., & Marsden, T. (2013). The influence of demand characteristics and social desirability on clients' ratings of the therapeutic alliance. *Journal of Clinical Psychology*, 69, 696–709. <http://dx.doi.org/10.1002/jclp.21946>
- Reese, R. J., Norsworthy, L. A., & Rowlands, S. R. (2009). Does a continuous feedback system improve psychotherapy outcome? *Psychotherapy: Theory, Research, & Practice*, 46, 418–431. <http://dx.doi.org/10.1037/a0017901>
- Reese, R. J., Toland, M. D., Slone, N. C., & Norsworthy, L. A. (2010). Effect of client feedback on couple psychotherapy outcomes. *Psychotherapy: Theory, Research, & Practice*, 47, 616–630. <http://dx.doi.org/10.1037/a0021182>
- Rise, M. B., Eriksen, L., Grimstad, H., & Steinsbekk, A. (2012). The short-term effect on alliance and satisfaction of using patient feedback scales in mental health out-patient treatment. A randomised controlled trial. *BMC Health Services Research*, 12, 348–360. <http://dx.doi.org/10.1186/1472-6963-12-348>
- Rise, M. B., Eriksen, L., Grimstad, H., & Steinsbekk, A. (2016). The long-term effect on mental health symptoms and patient activation of using patient feedback scales in mental health out-patient treatment. A randomised controlled trial. *Patient Education and Counseling*, 99, 164–168. <http://dx.doi.org/10.1016/j.pec.2015.07.016>
- Robert, C. (2007). *The Bayesian choice: From decision-theoretic foundations to computational implementation*. New York, NY: Springer Science & Business Media.
- Russo, J., Roy-Byrne, P., Jaffe, C., Ries, R., Dagadakis, C., Dwyer-O'Connor, E., & Reeder, D. (1997). The relationship of patient-administered outcome assessments to quality of life and physician ratings: Validity of the BASIS-32. *Journal of Mental Health Administration*, 24, 200–214. <http://dx.doi.org/10.1007/BF02898514>
- Safran, J. D., & Muran, J. C. (2000). *Negotiating the therapeutic alliance: A relational treatment guide*. New York, NY: Guilford Press.
- Safran, J. D., & Muran, J. C. (2006). Has the concept of the therapeutic alliance outlived its usefulness? *Psychotherapy: Theory, Research, & Practice*, 43, 286–291. <http://dx.doi.org/10.1037/0033-3204.43.3.286>
- Sapyta, J., Riemer, M., & Bickman, L. (2005). Feedback to clinicians: Theory, research, and practice. *Journal of Clinical Psychology*, 61, 145–153. <http://dx.doi.org/10.1002/jclp.20107>
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147–177.
- Schuman, D. L., Slone, N. C., Reese, R. J., & Duncan, B. (2015). Efficacy of client feedback in group psychotherapy with soldiers referred for substance abuse treatment. *Psychotherapy Research*, 25, 396–407. <http://dx.doi.org/10.1080/10503307.2014.900875>
- She, Z., Duncan, B. L., Reese, R. J., Sun, Q., Shi, Y., Jiang, G., . . . Clements, A. L. (2018). Client feedback in China: A randomized clinical trial in a college counseling center. *Journal of Counseling Psychology*. Advance online publication. <http://dx.doi.org/10.1037/cou0000300>
- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, M. S., . . . Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, 59, 22–33.
- Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology*, 78, 298–311. <http://dx.doi.org/10.1037/a0019247>
- Simon, W., Lambert, M. J., Harris, M. W., Busath, G., & Vazquez, A. (2012). Providing patient progress information and clinical support tools to therapists: Effects on patients at risk of treatment failure. *Psychotherapy Research*, 22, 638–647. <http://dx.doi.org/10.1080/10503307.2012.698918>
- Slone, N. C., Reese, R. J., Mathews-Duvall, S., & Kodet, J. (2015). Evaluating the efficacy of client feedback in group psychotherapy. *Group Dynamics*, 19, 122–136. <http://dx.doi.org/10.1037/gdn0000026>
- Snijders, T., & Bosker, R. (2012). *Multilevel analysis: An introduction to basic and applied multilevel modeling* (2nd ed.). London, UK: SAGE.
- Snyder, J., & Aafjes-van Doorn, K. (2016). Utilizing measure-based feedback in control-mastery theory: A clinical error. *Psychotherapy: Theory, Research, & Practice*, 53, 291–296. <http://dx.doi.org/10.1037/pst0000075>
- Sundet, R. (2012). Therapist perspectives on the use of feedback on process and outcome: Patient-focused research in practice. *Canadian Psychology*, 53, 122–130. <http://dx.doi.org/10.1037/a0027776>
- Tracey, T. J., & Kokotovic, A. M. (1989). Factor structure of the Working Alliance Inventory. *Psychological Assessment*, 1, 207–210. <http://dx.doi.org/10.1037/1040-3590.1.3.207>
- van Oenen, F. J., Schipper, S., Van, R., Schoevers, R., Visch, I., Peen, J., & Dekker, J. (2016). Feedback-informed treatment in emergency psychiatry: a randomised controlled trial. *BMC Psychiatry*, 16, 110–122. <http://dx.doi.org/10.1186/s12888-016-0811-z>
- Wampold, B. E. (2015). Routine outcome monitoring: Coming of age—With the usual developmental challenges. *Psychotherapy: Theory, Research, & Practice*, 52, 458–462. <http://dx.doi.org/10.1037/pst0000037>
- Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate* (2nd ed.). New York, NY: Routledge. <http://dx.doi.org/10.4324/9780203582015>
- Wampold, B. E., & Serlin, R. C. (2000). The consequence of ignoring a nested factor on measures of effect size in analysis of variance. *Psychological Methods*, 5, 425–433. <http://dx.doi.org/10.1037/1082-989X.5.4.425>
- World Health Organization. (2010). *International classification of diseases and related health problems* (10th ed.). Geneva, Switzerland: WHO press.
- Zilcha-Mano, S. (2017). Is the alliance really therapeutic? Revisiting this question in light of recent methodological advances. *American Psychologist*, 72, 311–325. <http://dx.doi.org/10.1037/a0040435>
- Zilcha-Mano, S., Dinger, U., McCarthy, K. S., & Barber, J. P. (2014). Does alliance predict symptoms throughout treatment, or is it the other way around? *Journal of Consulting and Clinical Psychology*, 82, 931–935. <http://dx.doi.org/10.1037/a0035141>
- 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. <http://dx.doi.org/10.1037/cou0000103>
- Zilcha-Mano, S., Lipsitz, I., & Errázuriz, P. (2018). When is it effective to focus on the alliance? Analysis of a within-client moderator. *Cognitive Therapy and Research*, 42, 159–171. <http://dx.doi.org/10.1007/s10608-017-9867-4>
- Zilcha-Mano, S., Muran, J. C., Hungr, C., Eubanks, C. F., Safran, J. D., & Winston, A. (2016). The relationship between alliance and outcome: Analysis of a two-person perspective on alliance and session outcome. *Journal of Consulting and Clinical Psychology*, 84, 484–496. <http://dx.doi.org/10.1037/ccp0000058>

Received April 16, 2018

Revision received September 10, 2018

Accepted September 13, 2018 ■