Do Treatment Manuals Undermine Youth–Therapist Alliance in Community Clinical Practice?

David A. Langer
Harvard University and Judge Baker Children’s Center

Bryce D. McLeod
Virginia Commonwealth University

John R. Weisz
Harvard University and Judge Baker Children’s Center

Author Note

David A. Langer, Department of Psychology, University of California, Los Angeles;
Bryce D. McLeod, Department of Psychology, Virginia Commonwealth University; John R. Weisz, Department of Psychology, Harvard University, and Judge Baker Children’s Center, Harvard Medical School.

David A. Langer is now at Department of Psychology, Harvard University, and Judge Baker Children’s Center, Harvard Medical School.

This research was supported in part by National Institute of Mental Health grants F31-MH079631 (DAL), F31-MH64993-01 (BDM), and R01-MH49522 (JRW), the John D. and Catherine T. MacArthur Foundation (Research Network on Youth Mental Health [JRW]), and the Norlien Foundation (JRW). Portions of these data were previously presented at the 2005 Association for the Advancement of Behavior Therapy, Washington, DC. We sincerely thank those who rated the many taped sessions in this study. We also thank the clinics, therapists, and families who participated in this research.
ALLIANCE AND MANUALIZED TREATMENTS

Address inquiries to David A. Langer, 53 Parker Hill Ave., Judge Baker Children’s Center, Boston, MA 02120. E-mail: dlanger@jbcc.harvard.edu
Abstract

Objective: Some critics of treatment manuals have argued that their use may undermine the quality of the client-therapist alliance. This notion was tested in the context of youth psychotherapy delivered by therapists in community clinics. Method: Seventy-six clinically-referred youths (57% female, aged 8–15, 34% Caucasian) were randomly assigned to receive non-manualized usual care or manual-guided treatment to address anxiety or depressive disorders. Treatment was provided in community clinics by clinic therapists randomly assigned to treatment condition. Youth–therapist alliance was measured using the Therapy Process Observational Coding System – Alliance (TPOCS-A) scale at four points throughout treatment, and the youth-report Therapeutic Alliance Scale for Children (TASC) at the end of treatment.

Results: Youths receiving manual-guided treatment had significantly higher observer-rated alliance than usual care youths early in treatment, the two groups converged over time, and mean observer-rated alliance did not differ by condition. Similarly, the manual-guided and usual care groups did not differ on youth-report of alliance. Conclusions: Our findings did not support the contention that using manuals to guide treatment harms the youth–therapist alliance. In fact, use of manuals was related to a stronger alliance in the early phase of treatment.

Keywords: Psychotherapy, Children, Adolescents, Alliance, Manualized Treatment
Do Treatment Manuals Undermine Youth–Therapist Alliance in Community Clinical Practice?

Treatment manuals play an important role in efforts to disseminate evidence-based treatments (EBTs) for youths. Manuals document treatment contents, guide the therapist, and support fidelity. Surveys and focus groups have revealed concerns that the use of treatment manuals may impede the development of a positive client–therapist alliance (e.g., Addis & Krasnow, 2000; Nelson, Steele, & Mize, 2006). However, concerns about treatments manuals may not be universal (Henggeler et al., 2007, 2008). These various perspectives highlight an important empirical question: Does the use of manual-guided treatment detract from alliance?

This study addresses that question.

Alliance, as typically defined, encompasses the affective and collaborative aspects of the client–therapist relationship (Elvins & Green, 2008). Alliance may play a critical role in both behavioral and non-behavioral treatments for youths since youths rarely self-refer and may resist engaging in treatment (Weisz, 2004). However, to our knowledge, no studies have investigated whether use of a treatment manual influences the alliance. A case can be made for influence in a negative direction; use of a manual may make the therapist appear rigid, or too driven by an agenda to attend to the youth. On the other hand, manual use might have a positive influence. Manuals tend to create a clear agenda and activities for treatment, and this may clarify the purpose and procedures for youths. These features, plus the structure and predictability of manual-guided treatment, may inspire confidence and trust in one’s therapist in ways that enhance alliance. So, the question of whether manuals influence the alliance persists.

To measure alliance, two strategies were considered: obtaining youth self-reports and observing and coding treatment sessions. Self-report taps youths’ perceptions of alliance and facilitates comparison of findings with extant literature (Elvins & Green, 2008; Shirk & Saiz,
Observational methods are not subject to demand characteristics and youths’ varying levels of ability to observe and report on their feelings (Shirk & Karver, 2003; Weisz, 2004). Given each method’s advantages, both methods were used in the present study.

Alliance was assessed among internalizing youths (i.e., those being treated for anxiety or depressive disorders) who had been randomized to manual-guided treatment (MG) or usual care (UC) in community mental health clinics. Externalizing youths are important, but the most common manualized interventions for such youths involve parent-training or medication (see Weisz & Kazdin, 2010), and neither context is ideal for assessing youth alliance. The most common manualized treatments for youth anxiety and youth depression involve cognitive-behavioral therapy (CBT; Weisz & Kazdin, 2010); CBT manuals for anxiety and for depression were used in this study. To maximize external validity, our sample included only clinically-referred youths treated in community clinics by therapists employed by the clinics. To increase internal validity, therapists were randomly assigned to use a treatment manual or to carry out their own preferred forms of UC with no treatment manual.

**Method**

**Participants**

**Youths.** Participants were 76 youths from seven community mental health clinics participating in the Youth Anxiety and Depression Study (see Southam-Gerow et al., 2010; Weisz et al., 2009), which assessed the effectiveness of manualized treatments relative to UC. Participants met DSM-IV criteria for an anxiety or depressive disorder that clinic staff, project staff, and parents agreed warranted primary treatment focus. The 76 youths (33 boys, 43 girls) were aged 8-15 ($M = 11.27$, $SD = 2.15$). Reflecting the area’s diversity, 34.21% were Caucasian, 32.89% Latino, 15.79% African American, 9.21% other (e.g., Native American), and 7.90% not
reported. Youths with mental retardation or psychotic symptoms were excluded. For the present study, inclusion criteria were: (a) attended a minimum of four sessions, at least three of which were recorded, and (b) had one therapist (cases with multiple therapists raise questions of how to model alliance over the course of treatment). Randomization resulted in 40 MG youths (five youths were excluded due to having fewer than four sessions) and 36 UC youths (three youths were excluded due to having fewer than four sessions). When a youth was randomly assigned to a condition, the youth was placed with a therapist who had already been randomized to that condition.

**Therapists.** The 59 therapists (mean age: 32.66 years, $SD = 8.72$; 78.00% female) were 18.64% social workers, 55.93% masters level psychologists, 13.56% doctoral-level psychologists, and 11.87% had other training backgrounds. Primary orientation was 30.36% psychodynamic, 23.21% CBT, 25.00% family systems, 5.36% eclectic, and 16.07% other.

**Assessment procedure.** Youths and their parents were interviewed at pre- and post-treatment. At each assessment, parents and youths provided written consent/assent and completed symptom measures separately. The study was IRB approved.

**Treatment Conditions**

**Manual group condition.** The Primary and Secondary Control Enhancement Training (PASCET; see Weisz et al., 2009) program for depression and the Coping Cat (CC; see Kendall et al., 1997) program for anxiety are individual-based manual-guided CBT programs. After training in PASCET or CC by study investigators, therapists used the respective manuals and were provided with weekly supervision. Mean MG treatment length was 17.78 sessions ($SD = 4.88$) and PASCET and CC did not differ significantly, $t(38) = 1.60, p = .12$. Cases were randomly selected for adherence coding from CC ($n = 10$ out of 15 total CC cases) and PASCET.
(n = 15 out of 25 PASCET cases). Expert raters coded all available tapes for each case; adherence to both manuals was very high. For details about the measures and methods used to assess adherence see Southam-Gerow et al. (2010) and Weisz et al. (2009).

**UC condition.** UC therapists did not use treatment manuals. Instead, they were instructed to employ the therapeutic procedures they used regularly and believed to be effective. No effort was made to influence UC in any fashion. The average number of sessions was 20.97 (SD = 14.89), not significantly different from MG, t(41.74) = 1.230, p = .226.

**Measures**

**Alliance measures.** The *Therapeutic Alliance Scale for Children* (TASC; Shirk & Saiz, 1992) is a 7-item youth-report alliance measure that was administered at post-treatment. The TASC has demonstrated acceptable psychometric properties in previous studies (Hawley & Weisz, 2005) as well as in the present study (α = .84). The *Therapy Process Observational Coding System for Child Psychotherapy – Alliance scale* (TPOCS-A; McLeod & Weisz, 2005) is a 9-item observer-rated alliance measure rated on a 6-point scale (0 = *not at all*, 5 = *great deal*). The TPOCS-A has demonstrated good reliability, internal consistency, and convergent validity (see McLeod & Weisz, 2005). In the present study, internal consistency was strong (α = .91), as was inter-rater reliability (ICC(1, 9) = .80).

**Other measures.** Youth symptomatology was assessed via the *Child Behavior Checklist* (CBCL; Achenbach, 1991).

**TPOCS-A Coders, Training, Coding, and Sampling Procedures.**

Training of the nine TPOCS-A coders (4 male; one licensed clinical psychologist, 4 graduate students, and 4 post-BAs) included reading the coding manual, reviewing specific session segments, and practice coding of sessions. Tapes were randomly assigned to coders.
Regular reliability assessments were performed and discussed in regular meetings to prevent rater drift. Four sessions were randomly selected from each case. To sample different therapy phases, each case was divided into early, middle, and late phases by dividing the total number of sessions by three. One session was randomly selected from “early” therapy (excluding the first session), one from “late” therapy (excluding the last session), and two from the “middle” (see Table 1 for the composition of each third). In cases where four taped sessions were not available (n=16), all available tapes were coded. Coders, who were uninformed as to any hypotheses or planned analyses, rated entire sessions; each session was double-coded.

Results
Youth and therapist pre-treatment differences between conditions were examined first and no significant differences were found (see Table 2). Next, potential correlates of alliance were examined, including youth age (child vs. adolescent), therapist age, youth gender, therapist gender, youth–therapist gender match, youth ethnicity, therapist ethnicity, and youth–therapist ethnicity match. Only variables related to youth gender were significant, so these variables are included in the results presented below, with the addition of therapist gender, a related construct.

For our main analyses, the first step focused on evaluating whether a difference existed on the child-report TASC. Post-treatment TASC scores for youths in the MG group (M = 23.03; SD = 4.65) were essentially identical to those in the UC group (M = 23.04; SD = 4.40), t(59) = 0.005 p = .996, d = 0.00). This test had power of .70 to detect an effect size of d = 0.58 (a medium effect) with a two-tailed α = .05. Overall observed alliance was not significantly different between conditions, γ1 = 0.20, p = .179 (0.26 SDs on the TPOCS-A). This test had a
power of .70 to detect a coefficient of 0.37 (0.47 SDs on the TPOCS-A scale; Snijders & Bosker, 1999). TASC scores and mean TPOCS-A scores were moderately correlated, $r = .481$, $p < .001$.

Next, MG and UC conditions were compared on observed alliance using multilevel models, modeling the progression of TPOCS-A alliance throughout treatment (see Table 3). All models were analyzed with HLM Version 6.08 (Raudenbush, Bryk, & Congdon, 2004). Time by itself did not significantly predict alliance scores (Model 1a). However, condition had a significant impact on alliance scores when included as a fixed factor on the model intercept and the slope of time (Model 1b). MG youths showed significantly higher alliance than UC youths in early treatment; as treatment continued alliance changed at a significantly different rate for MG than UC youths, and the two groups converged over time (see Figure 1). MG and UC youths did not differ significantly on alliance in middle or late treatment. Model 1c built on Model 1b by including youth and therapist gender. Treatment condition remained a significant predictor of the intercept and slope of alliance. Youth gender was a significant predictor of initial alliance levels, with females showing significantly higher alliance early in treatment.\(^2\) Power was .70 (two-tailed $\alpha = .05$) to detect a condition effect of approximately 0.50 $SDs$ on the intercept and 0.27 $SDs$ on the slope, on the TPOCS-A scale (Snijders & Bosker, 1999).

**Discussion**

Some critics have raised concerns that treatment manuals may negatively impact the quality of the alliance. The present study tested this possibility in the context of therapy for internalizing disorders, comparing alliance in youths who had been randomly assigned to manual-guided treatment versus UC without manuals. Using observational assessment to track alliance over time, results indicated that early in treatment alliance was stronger in the manual group than in UC. However, the groups converged over time, such that there was no significant
alliance difference between the two groups at mid-treatment or late in treatment; this was consistent with findings from the self-report alliance measure (TASC) showing no alliance difference at post-treatment. Taken together, the findings do not support the view that manuals undermine the alliance. Instead, the findings suggest that manual use may be associated with a stronger alliance early in treatment, and that alliance levels in manualized and non-manualized treatment may converge as treatment progresses. In none of our comparisons using observer- or youth-report alliance measures was manual use associated with a lower level of alliance than UC.³

One possible interpretation of these findings is that using treatment manuals may enhance alliance early in therapy. Perhaps youths respond well to the early structure and clear treatment agendas of manualized treatments. The information thus provided may be helpful to youths who are unfamiliar with therapy or uncertain about its purpose. Moreover, the manual may give the therapist and youth a specific agenda and thus an opportunity to discuss the ways in which desired outcomes will be achieved. Conversely, it is possible that alliance was stronger for MG early in treatment because the structured approach of the CBT manuals led to early symptom reduction. This possibility could not, however, be tested because outcomes were not assessed between pre- and post-treatment.

The convergence of MG and UC groups over time in treatment warrants attention. If higher alliance in the MG group early in treatment partly reflected the enhanced clarity of manualized treatment procedures, then group differences on that dimension might diminish over time as UC youths get to know their therapists and learn what to expect in therapy. This might well lead to increased alliance over time in the UC condition. By contrast, the structure and expectations (e.g., role plays, homework assignments) associated with manuals might lose some of their appeal over time, given their association with the “work” of therapy. More broadly, it is possible that
convergence over time could reflect an increasing impact of therapists’ individual characteristics and styles, as they become better acquainted with their clients, and an associated reduction in the salience and impact of the treatment procedures, manualized or not. These possibilities—all speculative at this point—would be best examined in future research that includes detailed assessment of clients’ perspectives on their therapists and therapy.

Even after treatment condition was taken into account, there was substantial variance around the alliance means, implying that factors other than treatment condition may relate to alliance. Of the factors tested, only youth gender was significantly related to alliance, yet there was insufficient power to test multiple correlates. Girls began therapy with significantly higher levels of alliance than boys. Girls may have been more ready to engage in treatment, or given the preponderance of female therapists it may have been easier for youths to form an alliance with same-gendered therapists. When the interaction was included in the model to test this hypothesis, the main effect of youth gender was no longer significant. The interaction of youth and therapist gender, which was significant only on the slope of alliance, $\gamma_{14} = -0.36, p = .04$, is difficult to interpret due to the high proportion of female therapists in the sample (78.00% female).

The limitations of the study suggest directions for future research. Observational coding of sessions is labor-intensive, so our sample of 76 represents significant effort, but robustness of findings can only be established through additional research. Our findings on correlates of alliance might change with larger samples. Self-reported alliance was only collected at post-treatment, but observational scores throughout treatment addressed that limitation. And, naturally, while UC therapists were unconstrained in their treatment provision, what usual care looks like may differ across therapists and settings. Our focus on youth–therapist alliance could be seen as a limitation. Parent–therapist alliance has been show to relate to symptom reduction,
attendance and persistence in therapy (Hawley & Weisz, 2005), so it also warrants study in the future. Similarly, the alliance–outcome relation in treatment delivered in practice settings is an important topic not addressed in this study but worthy of attention in the future given findings emphasizing the strength of this relation in real world settings (Hawley & Weisz, 2005; McLeod & Weisz, 2005; Shirk, Gudmundsen, Kaplinski, & McMakin, 2008). Lastly, generating a sampling design for treatments that vary in length presents a number of challenges. Our approach divided treatment into thirds based on the total number of sessions for each case. Although this sampling approach ensures that a session from each treatment phase was coded for each case, it also creates variability in the range of sessions included within each treatment phase. The mean session number in each third of treatment was comparable between conditions (see Table 1). However, the variability of the range of sessions within each treatment third must be considered when interpreting the study’s results.

Our findings suggest that the use of treatment manuals was associated with a stronger alliance early in therapy for youths with internalizing disorders, compared to non-manualized UC, and that there were no significant differences in alliance overall. A strong alliance early in treatment may be particularly important for youth therapy. Youths rarely refer themselves to treatment (Weisz, 2004), and even when they agree that they need help, youths, parents, and therapists often do not agree on referral problems (Hawley & Weisz, 2003). Thus, a critical task for therapists is to build enough of an alliance early in treatment to engage youths in the treatment process (Shirk & Saiz, 1992). Our findings suggest that using treatment manuals does not hinder this process, and may in fact be helpful in the task of early alliance building.
References


Footnotes

1 Prior to combining the protocols in our study, differences between treatment protocols on TASC scores and overall TPOCS-A scores were examined and found to be non-significant.

2 The youth gender x treatment condition interaction was also tested in a separate model and found to be nonsignificant on the intercept and slope of alliance.

3 The sample size does not provide sufficient power to test for small differences between the groups. However, it is worth noting that the MG group showed higher levels of alliance across treatment. Thus, the direction of the findings suggests that the use of treatment manuals does not negatively impact the alliance. For our main tests of interest – differences between conditions on the intercept and slope – the results were significant, attenuating concerns about power.
### Table 1

*Distribution of Sessions within Treatment Thirds between Conditions*

<table>
<thead>
<tr>
<th></th>
<th>MG</th>
<th>UC</th>
<th>Statistical Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 40)</td>
<td>(n = 36)</td>
<td></td>
</tr>
<tr>
<td><strong>Early Therapy</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum –</td>
<td>2 – 8</td>
<td>2 – 12</td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>3.89 (1.72)</td>
<td>4.00 (2.38)</td>
<td>t(73) = 0.22, p = .83</td>
</tr>
<tr>
<td>Median</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td><strong>Middle Therapy</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum –</td>
<td>2 – 18</td>
<td>3 – 26</td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>8.88 (3.20)</td>
<td>9.98 (7.07)</td>
<td>t(88.04) = 1.17, p = .25</td>
</tr>
<tr>
<td>Median</td>
<td>8</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td><strong>Late Therapy</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum –</td>
<td>4 – 28</td>
<td>6 – 49</td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>14.79 (4.38)</td>
<td>16.48 (11.29)</td>
<td>t(40.11) = 0.81, p = .42</td>
</tr>
<tr>
<td>Median</td>
<td>15</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* UC = Usual Care; MG = Manual-Guided.
Table 2

*Pre-treatment Differences between Manual-Guided and Usual Care Conditions*

<table>
<thead>
<tr>
<th>Variable</th>
<th>MG</th>
<th>UC</th>
<th>t (df) or ( \chi^2 )(df)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Youth</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>11.63 (2.33)</td>
<td>10.88 (1.88)</td>
<td>( t(74) = -1.55 )</td>
</tr>
<tr>
<td>CBCL Internalizing</td>
<td>67.95 (8.70)</td>
<td>68.70 (8.29)</td>
<td>( t(67) = 0.36 )</td>
</tr>
<tr>
<td>CBCL Externalizing</td>
<td>64.13 (9.15)</td>
<td>63.73 (10.62)</td>
<td>( t(67) = -0.17 )</td>
</tr>
<tr>
<td>CBCL</td>
<td>67.97 (8.90)</td>
<td>68.37 (8.64)</td>
<td>( t(67) = 0.18 )</td>
</tr>
<tr>
<td>Anxious/Depressed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBCL Total Problems</td>
<td>67.85 (7.39)</td>
<td>68.37 (8.02)</td>
<td>( t(67) = 0.28 )</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td>( \chi^2(1) = 1.21 )</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td>( \chi^2(5) = 4.84 )</td>
</tr>
<tr>
<td><strong>Therapist</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years of experience</td>
<td>9.40 (9.07)</td>
<td>7.71 (2.92)</td>
<td>( t(23.09) = -0.79 )</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td>( \chi^2(1) = 0.27 )</td>
</tr>
<tr>
<td>Degrees</td>
<td></td>
<td></td>
<td>( \chi^2(3) = 0.67 )</td>
</tr>
</tbody>
</table>

*Note.* UC = Usual Care; MG = Manual-Guided; CBCL = Child Behavior Checklist. CBCL scores were presented and analyzed as *t* values. All statistical tests were found to be non-significant (i.e., *ps > .10*).