

BRIEF REPORT

Psychotherapy for Depressed Youth in Poverty: Benchmarking Outcomes in a Public Behavioral Health Setting

Jonathan Kodet and Robert J. Reese
University of Kentucky

Barry L. Duncan
Better Outcomes Now, Jensen Beach, Florida

Robert T. Bohanske
Southwest Behavioral Health Services, Phoenix, Arizona

Research demonstrating the effectiveness of treatment with youth from low socioeconomic backgrounds is limited. To address this limitation, we evaluated pre–post psychotherapy treatment outcomes with youth presenting with depression-related diagnoses ($N = 469$) at a public behavioral health agency after they implemented a systematic client feedback approach as a quality improvement strategy. Clients were ethnically diverse youth at or under the poverty line. Treatment outcome was measured with the Outcome Rating Scale (Miller, Duncan, Brown, Sparks, & Claud, 2003) and the Child Outcome Rating Scale (Duncan, Sparks, Miller, Bohanske, & Claud, 2006). Benchmark methodology was used to compare effect size estimates to those achieved in randomized clinical trials. Average treatment effect sizes for the public behavioral health depression samples of children and adolescents ($d = 1.39$ and $d = 1.69$, respectively) were clinically superior to a waitlist benchmark drawn from clinical trials of youth depression, and clinically equivalent to a treatment benchmark drawn from youth depression clinical trials. Findings demonstrate that mental health services for depressed youth in poverty across an agency can be effective, and systematic client feedback may be a useful strategy to improve treatment outcomes.


Clinical Impact Statement

Question: Is psychotherapy utilizing systematic client feedback effective in reducing distress among depressed youth in poverty within a public behavioral setting? **Findings:** We found psychotherapy for depressed youth in a public behavioral setting in which systematic feedback was conducted showed similar effect sizes to treatments in clinical trials. **Meaning:** Systematic client feedback may be a useful quality improvement strategy for serving depressed youth. **Next Steps:** Although this study provides a more optimistic outlook on treatment of youth in a public behavioral setting than many previous studies, follow-up research is needed that uses a control condition to isolate the effects of client feedback to better understand how such processes benefit youth in psychotherapy.

Keywords: depression, poverty, youth, psychotherapy outcome, client feedback

Psychotherapy with youth has an established history of effectiveness. Meta-analyses and reviews of meta-analyses have found mostly small to moderate treatment effect size estimates (Klein, Jacobs, & Reinecke, 2007; Weisz, McCarty, & Valeri, 2006). When routine psychotherapy with youth in poverty is evaluated,

however, treatment outcomes are limited and typically less encouraging (e.g., $d = 0.25$, Farahmand et al., 2012; $d = 0.08$, Weersing & Weisz, 2002). Despite being most at risk for ongoing mental health problems (Reiss, 2013), youth from economically impoverished backgrounds often receive ineffective psychological treat-

 Jonathan Kodet and Robert J. Reese, Department of Educational, School, and Counseling Psychology, University of Kentucky; Barry L. Duncan, Better Outcomes Now, Jensen Beach, Florida; Robert T. Bohanske, Southwest Behavioral Health Services, Phoenix, Arizona.

Barry L. Duncan is a co-holder of the copyright of the Outcome Rating Scale/Session Rating Scale family of measures. The measures are free for individual use at <https://betteroutcomesnow.com>, but Barry L. Duncan receives royalties from licenses issued to groups and

organizations. In addition, there are web-based applications of the Partners for Change Outcome Management System, commercial products for which he receives royalties based on sales or profits from sales.

Correspondence concerning this article should be addressed to Jonathan Kodet, who is now at St. Cloud Hospital, CentraCare Health Plaza, 1900 CentraCare Circle, Suite 2375, St. Cloud, MN 56303. E-mail: jdko224@uky.edu

ment (Garland et al., 2013). Although the reasons for this finding have yet to be determined, possible explanations include inadequate and subpar services, youth and families being from more complex and disadvantaged backgrounds with compounding stressors, and inadequate access to resources (e.g., transportation).

For example, Weersing and Weisz (2002) used benchmarking methodology to compare treatment outcomes of ethnically diverse youth diagnosed with depression at six community mental health centers with outcomes derived from a meta-analysis of 13 clinical trials. Therapists reported using psychodynamic, cognitive, and behavioral interventions, with psychodynamic being the most commonly endorsed. At 3-month follow-up, the mean symptom severity on the self-report Children's Depression Inventory (Kovacs, 1992) in their treatment sample was almost identical to a benchmark no-treatment control group. When effectiveness studies are done in real-world settings, the results indicate a significant need for quality improvement.

Systematic client feedback offers one possible quality improvement strategy. It refers to the practice of monitoring client-reported outcome throughout treatment, which helps clinicians identify clients at risk for premature termination and to modify treatment with these clients. Solid evidence supports the use of systematic client feedback in psychotherapy with adults (Lambert, Whipple, & Kleinstäuber, 2018), but few studies have evaluated the benefit of client feedback with youth. Two studies (Bickman, Kelley, Breda, de Andrade, & Riemer, 2011; Nelson, Warren, Gleave, & Burlingame, 2013) both found that youth were shown to have faster rates of change when clinicians were provided with more frequent feedback. Lastly, Cooper, Stewart, Sparks, and Bunting (2013) found that school-based counseling incorporating systematic feedback was associated with large reductions in psychological distress for children ($d = 1.49$).

Although psychotherapy in a public behavioral health (PBH) setting that used a client feedback system, Partners for Change Outcome Management System (PCOMS; Duncan, 2011), was found to be effective (Reese, Duncan, Bohanske, Owen, & Minami, 2014), treatment that includes PCOMS has not been evaluated in a PBH setting with youth. The current study was designed to answer the question: In comparison with clinical trial benchmarks, is psychotherapy utilizing systematic client feedback effective in reducing overall psychological distress among youth with a depression-related diagnosis in a PBH setting? We focused on depression because it is one of the most common concerns addressed in psychotherapy for youth (Weisz et al., 2017) and provides an update to the discouraging Weersing and Weisz (2002) benchmark study focused on depression with youth.

Benchmarking Methodology

Evaluating the effectiveness of treatment as usual in clinical practice is challenging. Practice-based observational research typically does not allow for comparing treatment groups with a no-treatment control group, thus weakening internal validity. Benchmarking methodology was designed to address this and has been increasingly utilized in psychotherapy effectiveness studies (Lee, Horvath, & Hunsley, 2013; Reese et al., 2014) and consists of three steps: (a) construct pre-post benchmarks (i.e., ESs) from randomized clinical trials (RCTs) with waitlist control and intent-to-treat (ITT) samples (i.e., samples which include all client out-

comes irrespective of whether they completed treatment), (b) estimate the pre-post effect size (ES) of the naturalistic sample being evaluated, and (c) statistically compare the current sample ES against the constructed benchmarks derived from RCTs (Minami, Serlin, Wampold, Kircher, & Brown, 2008).

For statistical comparison, Serlin and Lapsley (1985) proposed a "good-enough principle" to allow for statistical testing with a range-null hypothesis to prevent rejection of a point-null hypothesis due to a large N . Recent benchmarking studies (Minami, Wampold, et al., 2008), have used an a priori margin of difference of 10%, indicating a clinically meaningful treatment effect (i.e., 90–110% of efficacy trial benchmark ESs). Because of this, a range-null hypothesis (e.g., $H_0: \delta_{PBHdep} \leq \delta_{ITT} - 10\%$) is used instead of a traditional point null hypothesis (e.g., $H_0: \delta_{PBHdep} = \delta_{ITT}$). Range-null hypotheses follow a noncentral t statistic (Serlin & Lapsley, 1985) and a normal distribution is approximated. Critical values are based on this range surrounding the benchmark ESs. Specifically, to determine if a naturalistic treatment ES is *equivalent* to treatment benchmark a critical value is calculated for the treatment group benchmark ES at $d_{ITT} - 10\%$, where $d_{ITT} - 10\%$ represents the lower bound of the 90–110% range. Thus, the null hypothesis is rejected if the difference exceeds the 90% lower bound of the benchmark ES. To determine if a naturalistic treatment ES is *superior* to a waitlist control benchmark then a critical value is calculated for the benchmark ES at $d_{WLC} + 10\%$ (or 110% of the benchmark). The naturalistic treatment ES must exceed this critical value.

For the current study, we used benchmarking methodology (Minami, Serlin, et al., 2008) to evaluate the effectiveness of psychotherapy services that used systematic client feedback with a sample of depressed youth (ages 6–17 years) who received individual treatment in a PBH setting. Our objectives were to examine if psychotherapy outcomes with the PBH youth sample were (a) clinically equivalent to treatment outcomes of a benchmark composed of RCT studies, and (b) superior to a benchmark of RCT studies that used a wait list control condition.

Method

Participants

Southwest Behavioral Health Services (SBHS) granted permission for data analysis from youth discharged cases between January 2008 and March 2014. SBHS is a not-for-profit comprehensive PBH organization serving a diverse range of individuals and families in Arizona. SBHS serves diverse youth at or below 100% of the federal poverty level, and clinicians use PCOMS comprehensively throughout its locations. Clients were excluded if intake scores were determined to be invalid (34 and above) according to parameters described in the PCOMS manual (Duncan, 2011). The final sample consisted of 469 youth who had a primary diagnosis of major depressive disorder, dysthymic disorder, depressed mood not otherwise specified (NOS), adjustment disorder with depressed mood, and adjustment disorder with mixed anxiety and depressed mood. All youth in the final sample engaged in individual therapy. The mean age was 12.88 years ($SD = 2.97$) with ages ranging from 6–17 years. Most clients were female (67.8%) with 29.4% identifying as White, 23.2% Latino/a, 4.1% African American, 2.3% "other," 1.9% as Native American, and 39.0% unknown. The

average number of sessions attended was 10.91, $SD = 10.26$, for adolescents ($n = 270$) and 12.71, $SD = 12.65$, for children ($n = 199$). Both treatment completers and noncompleters were in the sample. The number of clients who terminated prematurely (post-treatment scores above the clinic cutoff) were 81 (30%) adolescents and 79 (39.97%) children. All services for this sample were provided in English.

Therapists ($N = 86$) were predominantly female (84.2%) and White (88.1%), with 2.1% being African American and 9.8% Latino/a. Most (68.2%) had degrees in counseling, and the remaining had degrees in social work (12.7%), substance abuse counseling (11.3%), and psychology (8.8%). Therapists were all licensed and had at least a master's degree.

Measures

Outcome Rating Scale. The Outcome Rating Scale (ORS; Miller et al., 2003) is an ultrabrief four-item self-report outcome measure that uses a 10-cm visual analogue scale and is included in PCOMS (Duncan, 2011). The ORS items evaluate distress at four levels: Individual (personal), Interpersonal (intimate relationships), Social (work/school and relationships outside the home), and Overall. Clients make a hashmark on the visual analog scale to rate themselves on each of the items, which are scored to the nearest millimeter. Scores can range from 0–40 with lower scores indicating more distress. The ORS has a clinical cutoff score of 28 for clients aged 13–17 (Duncan et al., 2006). Scores less than 28 for adolescents indicate clinically significant psychological distress.

Psychometric evaluations of the ORS (Campbell & Hemsley, 2009) have shown the measure in related samples to have coefficient α s ranging from .87 to .93. Duncan et al. (2006) reported an α for youth aged 13–17 years as .93. Evidence for concurrent validity is suggested by a moderate bivariate correlation ($r = .53$) between the ORS and the Youth Outcome Questionnaire (Burlingame et al., 2001) with youth aged 13–17 years (Duncan et al., 2006).

Child Outcome Rating Scale. Like the ORS, the Child Outcome Rating Scale (CORS; Duncan et al., 2006) is an ultrabrief four-item self-report outcome measure that was designed for children aged 6–12. The same dimensions are evaluated, but the language is simplified, and “frowny” and “smiley” faces are used as anchors on similar 10-cm visual analog scales with scores 0–40. Based on 20,000 administrations from over 3,000 youth, coefficient α for the CORS was .84 (Duncan et al., 2006). Concurrent validity was tested with the Youth Outcome Questionnaire in a clinical sample of children 6–12 years resulting in a correlation of .61 (Duncan et al., 2006). The CORS has a clinical cutoff of 32, indicating a score less than 32 is typical for clinical populations for youth 6 to 12 years old (Duncan, 2014).

Procedures

SBHS implemented PCOMS in 2007. PCOMS uses the ORS or CORS to track outcome and facilitate discussions with clients regarding their treatment progress and the Session Rating Scale (SRS; Duncan et al., 2003) to monitor the therapeutic alliance. Monitoring each of these outcomes permits clinicians to intervene with clients who are not making adequate progress and discuss

adjustments to treatment needed in a collaborative manner. The ORS or CORS is completed and scored before each session and the SRS toward the end (e.g., last 5 min) of each session (the SRS data were not evaluated in this study). SBHS therapists received 12 hr of training on PCOMS (rationale for use; instruction on administration, scoring, and interpretation; and role plays with feedback) and then received annual 1-day booster trainings. Agency-wide policies enforced fidelity to PCOMS through routine supervisory meetings that include a protocol centered on the use of PCOMS data, and by requiring that therapists systematically collect PCOMS outcome data consistent with PCOMS procedures. For study inclusion, clients had to attend at least two sessions and have pre–post scores. We only included youth who had intake ORS/CORS scores below the clinical cutoff. These inclusion criteria were consistent with the RCT studies from Weisz et al. (2006) that we used as our benchmark.

Benchmark construction. The set of benchmarks includes an efficacy benchmark from the pre–post treatment outcomes of RCT treatment groups and a comparison benchmark from pre–post scores of the RCT waitlist/no-treatment control groups. We conducted a systematic search of the clinical trial literature from 2005 to August 2018 using inclusion criteria from recent benchmark studies (Reese et al., 2014). RCT studies had to include an ITT sample, as such samples are most comparable with effectiveness studies and utilize an outcome measure similar to the ORS/CORS (i.e., low reactivity and specificity). The search resulted in no RCTs meeting criteria with youth diagnosed with depression. Given this, we utilized Weisz et al.'s (2006) highly cited, rigorous meta-analysis of clinical trials to construct benchmarks for this study. First, we included 13 RCTs (of the total 35 studies) that reported information for ITT treatment groups in the ITT depression treatment efficacy benchmark. Second, we utilized 17 RCTs from Weisz et al. that reported means and standard deviations of waitlist/no treatment groups to calculate the waitlist control benchmark ES.

Depression efficacy trial benchmark effect size calculations. Next, the efficacy trial depression benchmarks were calculated using the procedures from recent benchmarking studies (Minami, Serlin, et al., 2008). We only included pre–post results from self-report outcome measures related to the primary diagnosis. When means and standard deviations were available for two self-report measures within a study, ESs were calculated separately and then aggregated using the mean of the ESs to obtain a single pre–post ES for the waitlist control group and ITT group. The formula ($d_i = [1 - (3/(4n - 5))] [(M_{\text{post}} - M_{\text{pre}})/SD_{\text{pre}}]$) for calculating an unbiased Cohen's d ES was used. After ESs for each study were calculated, they were aggregated across clinical trials to yield single ESs, serving as comparison benchmarks. The aggregation resulted in a waitlist control benchmark ES of $d_{\text{WL}} = 0.37$ and an ITT treatment group benchmark ES of $d_{\text{ITT}} = 1.01$.

Critical value calculation. The ES for the depression treatment condition has a corresponding critical value associated with the lower bound of the 10% range of clinical equivalence, and the waitlist condition ES has its corresponding critical value associated with the upper bound of the 10% range of clinical equivalence. The benchmarking hypotheses rely on a 95th percentile test statistic (e.g., $t_{(\text{ITT})\nu, \lambda; .95}$ and $t_{(\text{WL})\nu, \lambda; .95}$), which follows a non-central t distribution ($\nu = N - 1$ degrees of freedom) and has a noncentrality parameter $\lambda = \sqrt{N(d_{\text{ITT}} - 10\%)}$ or $\lambda = \sqrt{N(d_{\text{WL}} +$

10%). The critical values for the depression-related benchmarks are in Table 1.

PBH effect size calculations. Next, pre–post ESs (Cohen’s *d*) were calculated for the PBH depression sample using pre–post scores and standard deviations from client ORS/CORS scores. We used the same formula for the benchmarks to calculate ES estimates.

Results

The mean pre–post treatment ORS/CORS scores for the PBH depression sample were $M_{pre} = 20.75$ ($SD = 6.11$) and $M_{post} = 29.99$ ($SD = 7.58$), $d_{PBHdep} = 1.51$. The mean pre–post treatment ORS scores for adolescents were $M_{pre} = 19.53$ ($SD = 5.48$) and $M_{post} = 28.83$ ($SD = 7.50$), $d_{PBHdepADOL} = 1.69$. The mean pre–post treatment CORS scores for children were $M_{pre} = 22.41$ ($SD = 6.55$) and $M_{post} = 31.57$ ($SD = 7.68$), $d_{PBHdepCHILD} = 1.39$.

Compared against the ITT benchmark ($d_{ITT} = 1.01$) with a 10% a priori margin ($d_{ITT} [90\%] = 0.90$) and an associated critical value of $d_{cv(ITT)} = 1.00$, the observed PBH effect size estimate ($d_{PBHdep} = 1.51$) was statistically significant ($t = 32.70$, $df = 468$, $\lambda = 19.68$, $p < .001$). Specific to adolescents, the effect size estimate was statistically significant ($t = 27.77$, $df = 269$, $\lambda = 14.94$, $p < .001$). For children, the effect size estimate was also statistically significant ($t = 19.61$, $df = 198$, $\lambda = 12.82$, $p < .001$). All these analyses suggest at least clinical equivalence with ITT treatment efficacy from RCTs of youth in treatment for depression.

Compared against the waitlist control benchmark effect size estimate ($d_{WL} = 0.37$) given a 10% a priori margin ($d_{WL} [110\%] = 0.41$) and the associated critical value ($d_{cv(WL)} = 0.49$), the observed PBH effect size estimate ($d_{PBHdep} = 1.51$) was also statistically significant ($t = 32.70$, $df = 468$, $\lambda = 8.81$, $p < .001$), suggesting that treatment at PBH was clinically superior to waitlist control conditions from RCTs of youth in treatment for depression (see Table 1). When broken down by adolescents ($t = 27.77$, $df = 269$, $\lambda = 5.47$, $p < .001$) or children ($t = 19.61$, $df = 198$, $\lambda = 4.70$, $p < .001$), the findings are also statistically significant.

Discussion

To our knowledge, this study presents the first benchmarking analysis of treatment outcomes for youth with depression providing client feedback in psychotherapy at a PBH agency. Overall, we found that the treatment for the PBH youth sample was comparable with treatment effects found in youth depression RCTs. We

also found that that treatment effects for the PBH youth sample exceeded the waitlist control benchmark for youth.

Our results suggest that psychotherapy that includes a systematic client feedback system for youth provided in a large PBH setting is effective. The findings, though limited in generalizability, contrast with much of the existing literature on mental health treatment outcomes for financially disadvantaged youth. Previous studies have demonstrated that mental health care for youth from economically impoverished backgrounds results in negligible to small treatment effects (Brookman-Frazee et al., 2010; Farahmand et al., 2012). We know of only one previous benchmarking study of youth in poverty receiving mental health care in a PBH setting (Weersing & Weisz, 2002). Weersing and Weisz found that youth being treated for depression at a community mental health center on average had treatment outcomes equivalent to youth in waitlist/no-treatment control groups in RCTs. Their results dramatically differ from our findings, where youth at the PBH agency on average experienced large ES changes in treatment.

Several limitations need consideration when interpreting the findings in the present study. First, caution is warranted in interpreting the results, as we did not have a control condition to isolate the effects of PCOMS with the current sample. Rather, we provide evidence of effectiveness of psychotherapy with youth in poverty in a PBH setting. Findings in this setting may have been due to unknown therapist or supervisor effects or other treatment factors our design could not account for in this study. Because the data were collected within a single agency, perhaps other unknown characteristics of this specific setting accounted for the effectiveness of its youth program. Neither treatment approach nor therapist information in connection to each case was available in the current data set. Second, pre–post treatment analysis with only one outcome measure limits conclusions. Third, the lack of measure equivalency between the current sample and the benchmarks limits methodological rigor. Although the ORS/CORS are psychometrically acceptable, they are broad-based measures and yield less diagnostic specificity compared with the measures in the benchmarks. Alternatively, a case could be made that the lack of measure equivalency resulted in a more conservative comparison. Broad-based measures (i.e., low specificity) like the ORS/CORS often produce smaller effect sizes than high specificity measures (Lee, Jones, Goodman, & Heyman, 2005).

Although the current study cannot isolate effectiveness to the inclusion of client feedback, it is suggestive that it can be part of an effective treatment strategy for youth in poverty receiving psychotherapy services at a PBH agency. Specific to youth, the

Table 1
Effect Size Comparisons to Depression Benchmark Randomized Clinical Trial Studies

Sample	N	PBH <i>d</i>	ITT benchmark		Waitlist control benchmark	
			d_{cv}	<i>p</i>	d_{cv}	<i>p</i>
All youth	469	1.51	1.00	<.001	0.49	<.001
Age 13–17 (ORS)	270	1.69	1.03	<.001	0.51	<.001
Age 6–12 (CORS)	199	1.39	1.06	<.001	0.53	<.001

Note. PBH = public behavioral health; ITT = intent-to-treat; ORS = Outcome Rating Scale; CORS = Child Outcome Rating Scale; $d = [1 - (3/(4n - 5))] [M_{post} - M_{pre}/SD_{pre}]$; d_{cv} = critical effect size value required to attain statistical significance.

results are consistent with existing evidence that systematically including the child and adolescents' voice in clinical planning may improve the quality of mental health services and enhance treatment outcomes (Karver, Handelsman, Fields, & Bickman, 2006; McLeod, 2011).

More broadly, the current study has implications related to public policy for mental health treatment in the United States. Although millions of children in the United States are treated for mental health problems each year with an estimated 247 billion dollar cost (Centers for Disease Control and Prevention, 2013), most children receiving treatment through publicly funded agencies do not show clinical improvement (Warren, Nelson, Mondragon, Baldwin, & Burlingame, 2010). The current study results suggest this does not have to be the case.

Given the preliminary nature of the findings for the current study, we suggest further research evaluating client feedback as a quality improvement strategy in PBH agencies serving children and adolescents is needed. Specifically, we encourage clinical trial studies that use client feedback with youth in PBH settings. Our findings of large effect sizes and overall clinical equivalency with RCT treatment group outcomes support the contention that psychotherapy that includes systematic client feedback may be an effective quality improvement strategy in mental health services with youth in PBH—a sector of mental health where money and resources are limited.

References

- Bickman, L., Kelley, S. D., Breda, C., de Andrade, A. R., & Riemer, M. (2011). Effects of routine feedback to clinicians on mental health outcomes of youths: Results of a randomized trial. *Psychiatric Services, 62*, 1423–1429. <http://dx.doi.org/10.1176/appi.ps.002052011>
- Brookman-Frazee, L., Haine, R. A., Baker-Ericzen, M., Zoffness, R., & Garland, A. F. (2010). Factors associated with use of evidence-based practice strategies in usual care youth psychotherapy. *Administration and Policy in Mental Health, 37*, 254–269. <http://dx.doi.org/10.1007/s10488-009-0244-9>
- Burlingame, G. M., Mosier, J. I., Wells, M. G., Atkin, Q. G., Lambert, M. J., Whoolery, M., & Latkowski, M. (2001). Tracking the influence of mental health treatment: The development of the Youth Outcome Questionnaire. *Clinical Psychology and Psychotherapy, 8*, 361–379. <http://dx.doi.org/10.1002/cpp.315>
- Campbell, A., & Hemsley, S. (2009). Outcome Rating Scale and Session Rating Scale in psychological practice: Clinical utility of ultra-brief measures. *Clinical Psychologist, 13*, 1–9. <http://dx.doi.org/10.1080/13284200802676391>
- Centers for Disease Control and Prevention (CDC). (2013). Mental health surveillance among children - United States, 2005–2011. *Morbidity and Mortality Weekly Report, 62*, 1–35.
- Cooper, M., Stewart, D., Sparks, J., & Bunting, L. (2013). School-based counseling using systematic feedback: A cohort study evaluating outcomes and predictors of change. *Psychotherapy Research, 23*, 474–488. <http://dx.doi.org/10.1080/10503307.2012.735777>
- Duncan, B. (2011). *The Partners for Change Outcome Management System (PCOMS): Administration, scoring, interpreting update for the Outcome and Session Ratings Scale*. Jensen Beach, FL: Author.
- Duncan, B. L. (2014). *On becoming a better therapist: Evidence-based practice one client at a time* (2nd ed.). Baltimore, MD: American Psychological Association.
- Duncan, B., Miller, S., Sparks, J., Claud, D., Reynolds, L., Brown, J., & Johnson, L. (2003). The Session Rating Scale: Preliminary psychometric properties of a “working” alliance measure. *Journal of Brief Therapy, 3*, 3–12.
- Duncan, B., Sparks, J., Miller, S., Bohanske, R., & Claud, D. (2006). Giving youth a voice: A preliminary study of the reliability and validity of a brief outcome measure for children. *Journal of Brief Therapy, 5*, 71–87.
- Farahmand, F. K., Duffy, S. N., Taylor, M. A., DuBois, D. L., Lyon, A. L., Grant, K. E., . . . Nathanson, A. M. (2012). Community-based mental health and behavioral programs for low-income urban youth: A meta-analytic review. *Clinical Psychology: Science and Practice, 19*, 195–215. <http://dx.doi.org/10.1111/j.1468-2850.2012.01283.x>
- Garland, A. F., Haine-Schlagel, R., Brookman-Frazee, L., Baker-Ericzen, M., Trask, E., & Fawley-King, K. (2013). Improving community-based mental health care for children: Translating knowledge into action. *Administration and Policy in Mental Health and Mental Health Services Research, 40*, 6–22. <http://dx.doi.org/10.1007/s10488-012-0450-8>
- Karver, M. S., Handelsman, J. B., Fields, S., & Bickman, L. (2006). Meta-analysis of therapeutic relationship variables in youth and family therapy: The evidence for different relationship variables in the child and adolescent treatment outcome literature. *Clinical Psychology Review, 26*, 50–65. <http://dx.doi.org/10.1016/j.cpr.2005.09.001>
- Klein, J. B., Jacobs, R. H., & Reinecke, M. A. (2007). Cognitive-behavioral therapy for adolescent depression: A meta-analytic investigation of changes in effect-size estimates. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*, 1403–1413. <http://dx.doi.org/10.1097/chi.0b013e3180592aaa>
- Kovacs, M. (1992). *Children's Depression Inventory (CDI)*. Toronto, ON: Multi-Health Systems Inc.
- Lambert, M. J., Whipple, J. L., & Kleinstäuber, M. (2018). Collecting and delivering progress feedback: A meta-analysis of routine outcome monitoring. *Psychotherapy, 55*, 520–537. <http://dx.doi.org/10.1037/ps0000167>
- Lee, C. M., Horvath, C., & Hunsley, J. (2013). Does it work in the real world? The effectiveness of treatments for psychological problems in children and adolescents. *Professional Psychology: Research and Practice, 44*, 81–88. <http://dx.doi.org/10.1037/a0031133>
- Lee, W., Jones, L., Goodman, R., & Heyman, I. (2005). Broad outcome measures may underestimate effectiveness: An instrument comparison study. *Child and Adolescent Mental Health, 10*, 143–144. <http://dx.doi.org/10.1111/j.1475-3588.2005.00350.x>
- McLeod, B. D. (2011). Relation of the alliance with outcomes in youth psychotherapy: A meta-analysis. *Clinical Psychology Review, 31*, 603–616. <http://dx.doi.org/10.1016/j.cpr.2011.02.001>
- Miller, S., Duncan, B., 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.
- Minami, T., Serlin, R. C., Wampold, B. E., Kircher, J. C., & Brown, G. S. (2008). Using clinical trials to benchmark effects produced in clinical practice. *Quality and Quantity: International Journal of Methodology, 42*, 513–525. <http://dx.doi.org/10.1007/s11135-006-9057-z>
- Minami, T., Wampold, B. E., Serlin, R. C., Hamilton, E. G., Brown, G. S., & Kircher, J. C. (2008). Benchmarking the effectiveness of psychotherapy treatment for adult depression in a managed care environment: A preliminary study. *Journal of Consulting and Clinical Psychology, 76*, 116–124. <http://dx.doi.org/10.1037/0022-006X.76.1.116>
- Nelson, P. L., Warren, J. S., Gleave, R. L., & Burlingame, G. M. (2013). Youth psychotherapy change trajectories and early warning system accuracy in a managed care setting. *Journal of Clinical Psychology, 69*, 880–895. <http://dx.doi.org/10.1002/jclp.21963>
- Reese, R. J., Duncan, B. L., Bohanske, R. T., Owen, J. J., & Minami, T. (2014). Benchmarking outcomes in a public behavioral health setting: Feedback as a quality improvement strategy. *Journal of Consulting and Clinical Psychology, 82*, 731–742. <http://dx.doi.org/10.1037/a0036915>

- Reiss, F. (2013). Socioeconomic inequalities and mental health problems in children and adolescents: A systematic review. *Social Science and Medicine*, 90, 24–31. <http://dx.doi.org/10.1016/j.socscimed.2013.04.026>
- Serlin, R. C., & Lapsley, D. K. (1985). Rationality in psychological research: The good-enough principle. *American Psychologist*, 40, 73–83. <http://dx.doi.org/10.1037/0003-066X.40.1.73>
- Warren, J. S., Nelson, P. L., Mondragon, S. A., Baldwin, S. A., & Burlingame, G. M. (2010). Youth psychotherapy change trajectories and outcomes in usual care: Community mental health versus managed care settings. *Journal of Consulting and Clinical Psychology*, 78, 144–155. <http://dx.doi.org/10.1037/a0018544>
- Weersing, V. R., & Weisz, J. R. (2002). Community clinic treatment of depressed youth: Benchmarking usual care against CBT clinical trials. *Journal of Consulting and Clinical Psychology*, 70, 299–310. <http://dx.doi.org/10.1037/0022-006X.70.2.299>
- Weisz, J. R., McCarty, C. A., & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents: A meta-analysis. *Psychological Bulletin*, 132, 132–149. <http://dx.doi.org/10.1037/0033-2909.132.1.132>
- Weisz, J. R., Kuppens, S., Ng, M. Y., Eckshtain, D., Ugueto, A. M., Vaughn-Coaxum, R., . . . Weersing, V. R. (2017). What five decades of research tells us about the effects of youth psychological therapy: A multilevel meta-analysis and implications for science and practice. *American Psychologist*, 72, 79–117. <http://dx.doi.org/10.1037/a0040360>

Received September 2, 2018

Revision received April 16, 2019

Accepted April 17, 2019 ■

Members of Underrepresented Groups: Reviewers for Journal Manuscripts Wanted

If you are interested in reviewing manuscripts for APA journals, the APA Publications and Communications Board would like to invite your participation. Manuscript reviewers are vital to the publications process. As a reviewer, you would gain valuable experience in publishing. The P&C Board is particularly interested in encouraging members of underrepresented groups to participate more in this process.

If you are interested in reviewing manuscripts, please write APA Journals at Reviewers@apa.org. Please note the following important points:

- To be selected as a reviewer, you must have published articles in peer-reviewed journals. The experience of publishing provides a reviewer with the basis for preparing a thorough, objective review.
- To be selected, it is critical to be a regular reader of the five to six empirical journals that are most central to the area or journal for which you would like to review. Current knowledge of recently published research provides a reviewer with the knowledge base to evaluate a new submission within the context of existing research.
- To select the appropriate reviewers for each manuscript, the editor needs detailed information. Please include with your letter your vita. In the letter, please identify which APA journal(s) you are interested in, and describe your area of expertise. Be as specific as possible. For example, “social psychology” is not sufficient—you would need to specify “social cognition” or “attitude change” as well.
- Reviewing a manuscript takes time (1–4 hours per manuscript reviewed). If you are selected to review a manuscript, be prepared to invest the necessary time to evaluate the manuscript thoroughly.

APA now has an online video course that provides guidance in reviewing manuscripts. To learn more about the course and to access the video, visit <http://www.apa.org/pubs/journals/resources/review-manuscript-ce-video.aspx>.