



Published in final edited form as:

Adm Policy Ment Health. 2014 January ; 41(1): . doi:10.1007/s10488-012-0442-8.

Therapist Self-Report of Evidence-Based Practices in Usual Care for Adolescent Behavior Problems: Factor and Construct Validity

Aaron Hogue,

Division, National Center on Addiction and Substance Abuse (CASA) at Columbia University, 633 Third Avenue, 19th floor, New York, NY 10017, USA

Sarah Dauber, and

Division, National Center on Addiction and Substance Abuse (CASA) at Columbia University, 633 Third Avenue, 19th floor, New York, NY 10017, USA

Craig E. Henderson

Department of Psychology, Sam Houston State University, Huntsville, TX, USA

Abstract

This study introduces a therapist-report measure of evidence-based practices for adolescent conduct and substance use problems. The Inventory of Therapy Techniques—Adolescent Behavior Problems (ITT-ABP) is a post-session measure of 27 techniques representing four approaches: cognitive-behavioral therapy (CBT), family therapy (FT), motivational interviewing (MI), and drug counseling (DC). A total of 822 protocols were collected from 32 therapists treating 71 adolescents in six usual care sites. Factor analyses identified three clinically coherent scales with strong internal consistency across the full sample: FT (8 items; $\alpha = .79$), MI/CBT (8 items; $\alpha = .87$), and DC (9 items, $\alpha = .90$). The scales discriminated between therapists working in a family-oriented site versus other sites and showed moderate convergent validity with therapist reports of allegiance and skill in each approach. The ITT-ABP holds promise as a cost-efficient quality assurance tool for supporting high-fidelity delivery of evidence-based practices in usual care.

Keywords

Therapist-report fidelity; Adolescent behavior problems; Usual care; Family therapy; Cognitive-behavioral therapy; Motivational interviewing; Drug counselin

The emerging discipline of dissemination and implementation science is focused on discovering efficient methods for transplanting empirically supported treatments (ESTs) into routine practice settings, where they can take root as sustainable evidence-based practices (EBPs). There is now consensus that EST diffusion into usual care cannot succeed without rigorous implementation strategies designed to ensure that interventions are delivered to the appropriate populations and with adequate levels of fidelity to the main principles and procedures of the given model (Aarons et al. 2011; Damschroder and Hagedorn 2011). In laboratory settings such implementation strategies are usually called integrity or fidelity

methods, owing to their tight focus on the delivery of the intervention package itself in the context of controlled trials in which therapist, client, and organizational variables are favorable by design. In field settings implementation strategies are called quality assurance (QA) procedures (Bond et al. 2011; Schoenwald 2011). In addition to centerpiece fidelity methods for monitoring therapist performance, QA procedures also provide guidelines for establishing organizational characteristics and client assessment practices that promote sustained intervention fidelity and positive clinical outcomes (McHugh and Barlow 2010). Some QA systems also provide systematic feedback to providers and clients about fidelity data and client progress in an effort to enhance clinical judgment and effectiveness (Garland et al. 2010a, b).

This study introduces a therapist-report measure of fidelity to evidence-based interventions for adolescent behavior problems intended to serve as a QA tool in routine care. EBP fidelity measures of this kind are not novel; many EST developers have designed comprehensive provider training and QA procedures that are leased to consumers by purveyor organizations to facilitate model dissemination. These QA “superstructures” invariably contain several components (see Fixsen et al. 2005): (1) guidelines for selecting adoption-ready sites and identifying qualified staff for training; (2) standardized training toolkits that include a treatment manual, protocol for training workshops, demonstration videos and clinician workbooks, on-site supervision procedures, and fidelity checklists; (3) procedures for ongoing training and consultation from model experts, often including observational coaching of clinic cases based on audio or videotaped sessions; (4) continuous quality improvement procedures to evaluate implementation data collected on site, feed selected data back to therapists, and buttress organizational support; (5) certifications granted to providers who successfully complete training and maintain quality standards. A few EST developers have tested the impact of their QA superstructures on provider performance and client outcomes (e.g., Henggeler et al. 2008; Schoenwald et al. 2009), with promising results to date.

However, it is also true that importing QA superstructures from EST purveyors to support implementation in usual care requires often substantial changes in agency infrastructure, administrative and clinical supervision, material resources, and ongoing technical support (Hogue 2010). Whereas this level of resource commitment is obtainable by government-sponsored systems of care (Chambers et al. 2005; Zazzali et al. 2008), it is beyond reach for many behavioral health networks and stand-alone agencies. Moreover, there are few empirical guidelines that stipulate how many resources are needed to sponsor a “good enough” QA system, how much time and investment is required for imported QA procedures to become self-sustaining, or whether for more complex ESTs it is possible to graduate from outside technical support to localized QA.

For these reasons, developing efficient, cost-effective, and locally manageable QA tools to support high-fidelity EBP implementation in routine settings is a top priority for implementation science (Schoenwald et al. 2011). Such tools are not widely available within existing QA systems operated by managed care organizations; for the most part these systems employ metrics for quality behavioral care that draw from administrative databases and focus on pharmacological and inpatient care (Miranda et al. 2010). To oversee the selection and implementation of EBPs at the level of individual clients, new QA tools are needed that are psychometrically valid and focused on treatment fidelity: competent delivery of prescribed interventions and avoidance of proscribed interventions for target populations (Perepletchikova 2011). In addition, EBP assessment tools and designs intended for front-line clinics may require some compromises in rigor in order to achieve broader clinical relevance (Garland et al. 2010a, b; Kolko 2006): applicability to heterogeneous target clients with diverse profiles of co-occurring problems; ability to capture a wide variety of practices

rather than a narrow band of prespecified (manualized) techniques; and flexibility in marking the duration, targeted clients, ancillary interventions, and organizational features of treatment. Overall it appears that the most effective QA tools for assessing usual care interventions will (1) balance measurement precision with breadth of interventions assessed, (2) evaluate fidelity to both the content and process of EBP implementation, and (3) accommodate multiple uses and yield data that inform multiple levels and standards of accountability (Bearsley-Smith et al. 2008; Garland et al. 2010a, b; Schoenwald 2011).

The current study presents factor, discriminant, and convergent validity data for a therapist-report measure of EBPs delivered in usual care for adolescent conduct problems and substance use. Self-report measures of EBP fidelity offer several methodological strengths that may afford the desired balance between rigor and relevance in QA systems: they are quick, inexpensive, and non-intrusive; they capture the unique viewpoint of the provider delivering the interventions; and they can be completed throughout treatment, which facilitates measurement of infrequent but clinically meaningful interventions (Carroll et al. 1998; Weersing et al. 2002). If proven reliable and valid, therapist-report measures could support EBP implementation and inform quality improvement via feedback loops of several kinds: as a self-check by therapists to mark their own progress in treating individual cases; as a supervision aid for on-site and external trainers to monitor treatment fidelity; and as administrative data for stakeholders and external reviewers to evaluate therapist- and agency-level clinical performance (Bearsley-Smith et al. 2008; Bond et al. 2011; Garland et al. 2010a, b), among others.

This study introduces the Inventory of Therapy Techniques for Adolescent Behavior Problems (ITT-ABP), a measure of 27 intervention techniques representing four theoretical approaches to treating adolescents in outpatient behavioral care: cognitive-behavioral therapy (CBT), family therapy (FT), motivational interviewing (MI), and drug counseling (DC) (the derivation of ITT-ABP items is detailed in the Measures section; the items are listed in Table 1). These four approaches have a substantial base of empirical support for addressing adolescent conduct, delinquency, and/or substance use problems (for reviews see Becker and Curry 2008; Chorpita et al. 2011; Eyberg et al. 2008; see also Winters et al. 2007; Winters et al. 2000) and are widely endorsed in front-line settings. Moreover, the CBT, FT, and MI approaches each boast several manualized treatment models that have proven efficacious for a range of adolescent behavior problems (see Hogue and Liddle 2009; Miller and Rose 2009; Waldron and Turner 2008), making them ideal candidates to serve as transdiagnostic interventions capable of treating multiple disorders (Garland et al. 2010a, b; McHugh et al. 2009). The demand for transdiagnostic treatments is especially great in adolescent behavioral care, wherein co-occurring disorders and multisystem involvement are more rule than exception (Hawkins 2009; Kazak et al. 2010).

To date only a handful of studies have examined therapist-report fidelity measures designed to assess multiple treatment approaches commonly used in everyday practice. Reliability and construct validity data for such measures are generally encouraging, and as a group they cover a broad range of clinical problems and populations, including adult depression (Psychotherapy Practices Scale—Clinician Depression Care Version; Hepner et al. 2010), adult substance use (Clinical Practices Survey for Substance Use Disorders; Gifford et al. 2012), childhood disruptive behavior (Child Therapy Process Rating System; Hurlburt et al. 2010), and broad child psychotherapy approaches (Therapy Procedures Checklist; Weersing et al. 2002) (Treatment Recording Sheet; Bearsley-Smith et al. 2008).

The ITT-ABP is unique among therapist-report, multiapproach measures of usual care practices in several ways. It focuses on adolescents rather than children (vs. Hurlburt et al. 2010) and externalizing rather than internalizing problems (vs. Bearsley-Smith et al. 2008).

It is designed as a post-session measure rather than a monthly report (vs. Monthly Treatment and Progress Summary; Nakamura et al. 2007) or a post-termination measure (vs. Weersing et al. 2002) to maximize its acuity. And rather than focusing on session content (vs. Session Report Form; Kelley et al. 2010), it assesses discrete intervention techniques (aka clinical strategies, or practice elements), which is “middle ground” between molar versus molecular specification of treatment processes and thus well-suited for describing clinical practices in usual care (Garland et al. 2010a, b). Most importantly, ITT-ABP items derive from observational fidelity scales that were constructed during controlled trials of manualized treatments (see Methods). There is considerable enthusiasm for using observational fidelity measures of ESTs as a foundation for developing QA tools in routine care (Gifford et al. 2012; Schoenwald et al. 2011), on the premise that this method is most likely to yield instruments with both strong psychometric properties and clinical validity in measuring specific EBPs.

This study assessed the factor structure and construct validity of the ITT-ABP based on post-session self-report data from 32 community therapists treating 71 adolescents at six outpatient sites. Therapists and clients were participating in a randomized field trial testing the effectiveness of usual-care treatment for adolescent behavior problems. Clients were assigned to one of two study conditions: (a) Routine Family Therapy (RFT), consisting of a single site that featured family therapy as its routine standard of care for mental health treatment; or (b) Treatment As Usual (TAU), consisting of five sites in the same catchment area that featured treatment approaches other than family therapy. All study therapists operated under routine practice conditions without benefit of external training, supervision, or clinical resources from the research team.

The current study had two primary aims. First, the factor structure of the 27-item ITT-ABP was determined using a cross-validation method whereby factors were derived from half of the sample via exploratory factor analysis and then confirmed on the remaining half via confirmatory factor analysis. Based on factoring results, ITT-ABP scales were created and examined for internal consistency, intercorrelation patterns, and clinical coherence. Second, the construct validity of derived ITT-ABP scales was examined by testing their (1) discriminant validity in distinguishing among sites with different espoused treatment orientations and (2) convergent validity in the form of correlations with therapist-reported allegiance and skill in each of the four EBPs: CBT, FT, MI, and DC. It was predicted that factor analyses would reveal and confirm an underlying scale structure consisting of four discrete factors matching the original scales of the source observational measures: CBT, FT, MI, and DC. It was also predicted that the RFT condition would report greater use of techniques associated with the family therapy approach, and lesser use of techniques associated with the other approaches, compared to TAU. Finally, it was predicted that therapist self-reported levels of allegiance and skill in each of the four EBPs at baseline would predict their later reports of EBP utilization during sessions with study cases.

Method

Study Clients

Client participants ($N = 71$) included adolescents (44 % male; Mean age 15.4 years [$SD = 1.4$]) and their primary caregivers. Self-reported race/ethnicity was Hispanic (62 %), African American (17 %), multiracial (11 %), White (6 %), and other (4 %). Households were headed by single parents (62 %), two parents (21 %), grandparents (7 %), or other (3 %); 42 % earned less than \$15,000 per year, 14 % received public assistance, and 41 % reported a history of child welfare involvement. Adolescents were referred primarily from schools (78 %) but also from family service agencies (14 %) and other sources (8 %); 24 % were involved in the juvenile justice system at referral. Rates of psychiatric diagnosis based on

the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR; American Psychiatric Association 2000), with diagnoses given for meeting threshold based on either adolescent or caregiver report, were: Oppositional Defiant Disorder (ODD) = 90 %, Attention-Deficit/Hyperactivity Disorder = 75 %, Conduct Disorder (CD) = 48 %, Mood Disorder or Dysthymia = 44 %, Substance Use Disorder (SUD) = 23 % (69 % cannabis use, 31 % alcohol), Generalized Anxiety Disorder = 17 %, Posttraumatic Stress Disorder = 20 %. A total of 87 % of the sample was diagnosed with more than one disorder.

Client Recruitment and Randomization

Client participants were part of a randomized field trial designed to identify adolescents with untreated behavioral health problems, enroll them in available outpatient treatment services, and assess treatment effects up to one year later. Research staff developed a referral network of high schools, family service agencies, and youth programs serving clients in primarily inner-city areas of a large northeastern city. Network partners made referrals to research staff during site visits and also by phone and confidential email. Staff then contacted referred families by phone and offered them an opportunity to participate in a two-part baseline interview to assess the reason for study referral and discuss study enrollment.

After completion of the baseline interview, adolescents (1) who met diagnostic criteria for ODD, CD, and/or SUD and (2) whose families were interested in receiving outpatient treatment services, were randomly assigned to one of two study conditions: Routine Family Therapy or Treatment As Usual. Urn randomization procedures were employed to promote balance between conditions on three variables: sex, ethnicity, and juvenile justice involvement. One TAU site that specialized in addiction treatment was withheld from randomization of ODD and CD cases, and one TAU site that did not accept substance users was withheld from randomization of SUD cases. Significant differences between study conditions were found for one variable: TAU had a higher proportion of caregivers previously investigated by the child welfare system ($\psi^2(1) = 9.4, p < .01$). No other group differences on demographic or diagnostic variables were found.

Study Sites and Therapists

All six treatment sites were outpatient clinical settings that accepted study cases as standard community referrals. No external training, financial support, or logistical support of any kind was provided to treat study cases, and therapists were not required to alter their clinical practices in any way. Therefore, all study sites provided usual-care services to referred families. Each site prescribed weekly treatment sessions and had in-house psychiatric support. Sites were in close geographical proximity and easily accessible to all families via public transportation.

Routine Family Therapy (RFT)—The RFT condition consisted of a single community mental health clinic that featured family therapy as the standard-of-care approach for behavioral interventions with youth. RFT therapists ($n = 15$, who treated 38 cases total) were licensed Marriage and Family Therapists, licensed Social Workers with training in family therapy, or advanced clinical trainees with family therapy experience. All RFT therapists received routine in-house training and supervision designed to promote the signature clinical techniques of the FT approach; no specific manualized EST was being used. All site therapists who volunteered were accepted into the study. Participating therapists ranged in age from 28 to 59 years; 8 were female, 7 were Hispanic American, and as a group they averaged 3.1 years ($SD = 4.3$) postgraduate therapy experience.

Treatment As Usual (TAU)—This condition included a set of five clinics in order to capture a representative range of EBPs and sample the full spectrum of outpatient treatment

options available for adolescent behavior problems. Among the five TAU sites, two were community mental health clinics (8 therapists total) with organizational profiles that basically matched the RFT site. Two other sites were outpatient clinics (6 therapists total) housed within the child and adolescent psychiatry departments of teaching hospitals. The fifth site was an independent addictions treatment clinic (3 therapists) with an adolescent program that featured group-based treatment with supportive individual sessions. TAU sites provided treatment services that, according to focus group feedback (see Measures section), were nominally consistent with the MI, CBT, and/or DC approaches. No TAU site contained a supervisor or staff therapist with extensive training in family therapy, and no site appeared to promote or feature implementation of signature FT techniques. All therapists who volunteered at each site were accepted into the study. Across the five sites, participating therapists ($n = 17$, who treated 33 cases total) ranged in age from 25 to 45 years; 11 were female, 10 were European American, and as a group they averaged 3.2 years ($SD = 2.9$) postgraduate therapy experience.

Measures

Inventory of Therapy Techniques—Adolescent Behavior Problems (ITT-ABP)

—The 27-item ITT-ABP was designed as a QA tool for collecting post-session therapist-report data on implementation of discrete treatment techniques associated with the CBT, FT, MI, and DC approaches, respectively. ITT-ABP items were derived from four validated observational fidelity scales for manualized ESTs representing each approach, using a two-stage development process. First, the original psychometric studies of the four scales under consideration (see below) were reviewed to examine strength of factor loadings and interrater reliability for scale items. Second, two focus groups were conducted at each study site with all available therapists to review prospective ITT-ABP items, gather information about fit between prospective items and the treatment practices favored at each site as reported in the focus groups, and accordingly trim prospective items, to create the final set of four ITT-ABP theoretical scales: CBT, FT, MI, and DC. Table 1 lists all 27 items across all four theoretical scales. The ITT-ABP measures the thoroughness and frequency (which jointly reflect EBP quantity, as opposed to quality) with which each technique was utilized in a just-completed session, based on a 5-point Likert-type scale: 1 = *Not at all*, 2 = *A little bit*, 3 = *Moderately*, 4 = *Considerably*, 5 = *Extensively*.

ITT-ABP items representing the FT approach (7 items) and CBT approach (7 items) were drawn from the *Therapist Behavior Rating Scale (TBRS)* (Hogue et al. 1998), a macroanalytic observational tool designed to identify therapeutic techniques prescribed by FT and CBT for adolescent behavior problems. The TBRS has demonstrated strong psychometric properties in studies of treatment adherence (Diamond et al. 2007; Hogue et al. 2008a, b; Hogue et al. 1998), therapist competence (Hogue et al. 2008a, b), and fidelity-outcome links (Hogue et al. 2004; Hogue et al. 2006; Hogue et al. 2008a, b) with samples including drug-using, conduct-disordered, and depressed teens. Factor analyses of the TBRS revealed a FT dimension including family relationship and family interaction items, and a CBT dimension including behavior/skills items and cognition-focused items (Hogue et al. 1998, 2004). ITT-ABP items representing the MI approach (7 items) and DC approach (6 items) were drawn from the Motivational Enhancement Therapy and Twelve Step Facilitation subscales of the Yale Adherence and Competence Scale (YACS) (Carroll et al. 2000). The YACS is a general system for rating adherence and competence in delivering behavioral treatments for adult substance use disorders. It has demonstrated strong reliability and validity for MI and DC fidelity scores in an efficacy trial comparing CBT, 12-step facilitation, and clinical management for adult substance use disorders (Carroll et al. 2000) as well as multisite effectiveness trials comparing manualized MI treatment to usual care (Ball et al. 2007; Carroll et al. 2006). Although the YACS was validated on adult clinical

samples, it was chosen as the source measure for MI and DC items because no observational fidelity measures for these approaches had been validated on teenage samples at the time of this study, and during focus groups study therapists attested to the applicability of YACS source items for their usual practices with adolescent clients. All ITT-ABP items retained the content and molarity of the corresponding items from the source measures.

Therapist Self-Reported Proficiency in EBPs for ABP—We developed two face-valid questions to assess therapists' own judgments about their clinical orientation and technical skill related to the four evidence-based treatment approaches for adolescent behavior problems contained in the ITT-ABP: cognitive-behavioral therapy, family therapy, motivational interviewing, and drug counseling. Therapist were asked to self-rate their degree of allegiance to, as well as their skills in implementing, each of the four approaches using a 5-point scale: 1 = *None*, 2 = *A Little*, 3 = *Moderate*, 4 = *Considerable*, 5 = *High*. Thus each therapist provided eight responses (one allegiance and one skill rating each approach); the questions were administered prior to assigning study cases. Because allegiance and skill ratings were moderately to highly correlated within each treatment approach (CBT: $r = .58, p < .01$; FT: $r = .73, p < .001$; MI: $r = .75, p < .001$; DC: $r = .85, p < .001$), allegiance and skill scores were averaged to create a single "proficiency" score for each approach to minimize redundancy in study analyses.

Implementation Data Collection Procedures and Sampling Rates

Prior to seeing study cases, study therapists participated in a one-hour group session introducing the purpose of the ITT-ABP and briefly reviewing all items. At this meeting, therapists were given an ITT-ABP scoring manual (available upon request from the first author) that contained onepage descriptions of every item, including exemplar statements of what a therapist might say in session when implementing the given technique; item descriptions and exemplars were distilled from the corresponding item materials contained in the TBRS and YACS coding manuals. Therapists were then asked to complete a checklist after every treatment session with study clients. During the study, therapists were prompted to submit completed checklists for active cases on a weekly basis but received no further ITT-ABP training and no feedback on submitted data.

Of the 197 cases randomized into the trial at the time of final data collection for the current study, 80 (41 %) completed clinical intake and attended at least 1 treatment session at the assigned site. Of these 80 treatment attenders, 71 (89 %) were included in this study by virtue of having at least one submitted ITT-ABP checklist. Overall, study cases attended a total of 1156 sessions ($M = 16.3$; $SD = 16.8$; range = 1–104). Of these 1156 attended sessions, checklists were submitted for 822 sessions (71 %), including 505 checklists from 708 RFT sessions (71 %) and 317 checklists from 448 TAU sessions (71 %). The 15 therapists in RFT submitted 505 checklists across 38 cases, averaging 33.7 ($SD = 30.8$) checklists submitted per therapist (range = 2–122) and 13.2 ($SD = 13.3$) per case (range = 1–54). The 17 therapists in TAU submitted 317 checklists across 33 cases, averaging 18.6 ($SD = 11.2$) checklists submitted per therapist (range = 1–36) and 9.6 ($SD = 8.1$) per case (range = 1–25). There was no significant between-condition difference in number of checklists submitted.

Plan of Analysis

The factor structure of the ITT-ABP was tested using a two-step procedure. The sample was randomly split into two subsamples. First, principal components exploratory factor analysis (EFA) with direct oblimin rotation was conducted on the first half of the sample ($N = 410$) to determine the optimal number of factors. One, two, three, and four factor solutions were extracted, and the optimal solution was selected based on a combination of decline in

eigenvalues and interpretability. To make the final model solution as parsimonious as possible, we trimmed items with less than optimal factor loadings ($<.40$). Following EFA, confirmatory factor analysis (CFA) was conducted on the other half of the sample ($N = 412$) to evaluate model fit. As standard in CFA, we assessed model fit using the model Chi-square statistic and two supplementary fit indices, RMSEA and CFI. RMSEA values of .06 and below and CFI values of .95 and above indicate strong model fit, and CFI $\geq .90$ and RMSEA $\leq .08$ indicate adequate fit (Browne and Cudeck 1993; McDonald and Ho 2002). Both EFA and CFA were conducted in Mplus version 6.11 and used the sandwich variance estimator to account for the nested structure of the data (Asparouhov 2005). As described below in the Results section, three factors were derived from the ITT-ABP scale. Descriptive statistics and internal consistency reliability were examined for the factor scales separately for the RFT and TAU groups, as well as for three subgroups within the TAU group defined by treatment type: community mental health clinic (CMHC), addictions clinic, and outpatient psychiatry clinic.

Condition differences on the ITT-ABP scales were examined using multilevel mixed-effects linear regression models to account for the dependence of observations produced by the nested data structure (sessions nested within clients within therapists). Multilevel mixed-effects models use adaptive quadrature estimation to allow for the estimation of both fixed and random effects (Rabe-Hesketh and Skrondal 2005) and produces unbiased standard error estimates for nested data by accounting for correlated error terms (Hedeker et al. 1994; Raudenbush and Bryk 2002). Separate models were conducted to examine condition (RFT vs. TAU) effects on each derived ITT-ABP scale, controlling for client gender, age, ethnicity and study track (mental health case versus substance use case). An additional set of multilevel models were conducted to examine differences between RFT and the CMHCs within the TAU group on use of the ITT-ABP scales. Cohen's d effect sizes (interpretable as $.20 = \text{small}$, $.50 = \text{medium}$, $.80 = \text{large}$; Cohen 1988), along with the random effects coefficients for therapist and client, were calculated for model.

The final set of analyses examined the impact of therapist proficiency on use of the derived ITT-ABP scales. First, independent samples t -tests were conducted to compare RFT versus TAU therapists on proficiency within each treatment approach. Next, multilevel mixed-effects linear regressions were conducted to examine the effects of therapist proficiency on ITT-ABP scale scores. As described above, all models accounted for the nested data structure and controlled for client covariates (gender, age, ethnicity, track) and study condition (RFT vs. TAU). Four separate models were conducted to examine the predictive effects of FT proficiency on FT utilization, MI proficiency on MI/CBT utilization, CBT proficiency on MI/CBT utilization, and DC proficiency on DC utilization. All models examined main effects of proficiency score on ITT-ABP scale as well as interactions between proficiency score and study condition.

Results

Factor Validity: Factor Structure of the ITT-ABP

Exploratory factor analysis was conducted on half of the sample to determine the optimal factor structure of the scale. EFA was conducted using the Weighted Least Squares robust estimator, and the sandwich estimator to account for the nested structure of the data. Eigenvalues were 9.81 for one factor, 4.06 for two factors, 2.75 for three factors, and 1.60 for four factors. One-, two-, three-, and four-factor solutions were examined and the three-factor solution was selected as the most substantively interpretable. Table 1 contains the 27 ITT-ABP items organized in terms of their loadings on the three derived factors. For each item, the theoretical factor (CBT, FT, MI, or DC) associated with that item is also included in the table. Two items (Homework, Equality and collaboration) did not load above $.40$ on

any factor and were therefore dropped from further analyses. Correlations among the three factors were: Factor 1 and Factor 2: $r = .37$; Factor 1 and Factor 3: $r = .07$; Factor 2 and Factor 3: $r = .17$.

Factor 1, named Drug Counseling Interventions, included 9 items, 8 of which were from the DC theoretical scale and one from the CBT theoretical scale (Non-drug prosocial activities). Factor loadings ranged from .97 for “Explored positive and negative effects of substance use and the pros and cons of abstinence” to .44 for the CBT item “Helped client develop friends, relationships, and social activities that are non-drug related.” With the exception of this item, all other items loaded above .78 on this factor. Factor 2, named MI/CBT Interventions, included 8 items drawn from the MI and CBT theoretical scales. Factor loadings ranged from .81 for “Facilitated the client’s awareness of discrepancies between current problematic behaviors and future goals” to .52 for “Utilized behavioral interventions or taught relaxation exercises.” All items loaded above .50 on this scale. Factor 3, named Family Interventions, included 8 items, all but one (Sets agenda) from the FT theoretical scale. Factor loadings ranged from .73 for “Worked on enhancing communication and attachment between family members” to .46 for “Worked individually with adolescent or caregiver to prepare for family session.” Across the board, relatively low factor loadings (below .50) may be indicative of overall low therapist endorsement rates of the given interventions.

Following EFA, the three factor solution was confirmed on the remaining half of the sample using confirmatory factor analysis (CFA). The CFA model that was tested mirrored the three-factor EFA solution presented in Table 1. The two items that did not load $>.40$ on any factor in the EFA (Homework; Equality and Collaboration) were not included in the CFA model. Thus, the CFA model included three latent factors: Drug Counseling Interventions (measured by 9 observed indicators), MI/CBT Interventions (measured by 8 observed indicators), and Family Interventions (8 observed indicators). The initial CFA model that was tested fit the data well and was theoretically viable, so that no modifications were made and no constraints on the model were needed to improve fit. The three factors derived from the EFA were specified in the CFA model and model fit was evaluated by examining the Chisquare, RMSEA, CFI and TLI. The following fit indices were obtained for the CFA model: $\chi^2(272) = 388.011, p < .001$; RMSEA = .032 (90 % CI: .025-.039); CFI = .960; TLI = .956. Evaluation of the fit indices indicated that the three factors obtained in the EFA fit the data well. Bivariate correlations between the latent factors were: Drug Counseling and MI/CBT: $r = .52$, Family and Drug Counseling: $r = .33$, and Family and MI/CBT: $r = .48$.

Factor Validity: Internal Consistency of the ITT-ABP Scales

Scale scores for each of the three ITT-EBP factors were calculated by averaging the scores for all items loading on the corresponding factor. For the full sample ($N = 822$), average scores were 1.27 ($SD = .58$) on the Drug Counseling Interventions scale, 2.37 ($SD = .82$) on the MI/CBT Interventions scale, and 2.43 ($SD = .78$) on the Family Interventions scale. Cronbach’s alphas for the full sample were .90 for Drug Counseling, .87 for MI/CBT, and .79 for Family. Table 2 contains average scale scores and alphas for the RFT and TAU conditions separately, as well as for the subgroups of CMHCs, Addictions Clinic, and Outpatient Psychiatry Clinics within the TAU condition. As shown in the table, reliability was adequate for all three scales within both study conditions. The modest reliability for Drug Counseling in the RFT condition ($\alpha = .56$) was likely due to low endorsement rates of drug counseling interventions among family-oriented therapists. Within the TAU group, though the overall alpha for Drug Counseling was high ($\alpha = .95$), low reliability for this scale was found within CMHCs and Outpatient Psychiatry clinics, again due to low endorsement rates of Drug Counseling items among therapists at these sites.

Discriminant Validity: Between and Within Condition Comparisons of ITT-ABP Scales

As described above, multilevel mixed-effects linear regression models were conducted to test for between-condition differences in the three ITT-ABP scales, controlling for covariates. No significant differences between the RFT and TAU conditions were found on the Drug Counseling or MI/CBT scales. Significantly higher scores on the Family scale were found for the RFT condition compared to the TAU condition ($B = .53$, $SE = .19$, $z = 2.73$, $p < .01$, $d = .68$). After accounting for condition, a statistically significant amount of variability remained within therapists (random effect coefficient = $.46$, $SE = .09$) and clients (random effect coefficient = $.31$, $SE = .05$). In order to examine whether this pattern of findings would be replicated when limiting the TAU condition to just CMHCs (given that the single RFT site was a CMHC), we conducted another set of multilevel models comparing RFT to the two CMHC sites within TAU. As above, no differences were found for Drug Counseling or MI/CBT, and the RFT condition had higher Family scores than did TAU CMHCs ($B = .76$, $SE = .21$, $z = 3.61$, $p < .001$, $d = .97$). And again, there was significant variability within therapists (random effect coefficient = $.39$, $SE = .10$) and clients (random effect coefficient = $.32$, $SE = .06$). Inspection of the mean scale scores in Table 2 reveals that within RFT, Family Intervention items were endorsed most frequently, as expected. Within TAU, MI/CBT items were endorsed most frequently. This pattern was repeated within each of the TAU subgroups. The Addictions Clinic had the highest average score on Drug Counseling items, also consistent with expectations.

Convergent Validity: Relation of Therapist Proficiency Scores to ITT-ABP Scales

Independent samples *t*-tests were conducted to compare RFT and TAU therapists on their self-reported levels of proficiency (allegiance and skill) in each of the four treatment approaches (FT, MI, CBT, DC). Due to the small number of therapists with data available for these analyses (due to missing data, proficiency scores were available for only 10 RFT therapists and 16 TAU therapists), effect sizes are more interpretable than statistical significance. Results are presented in Table 3. As expected, RFT therapists reported a large effect size for greater proficiency in FT compared to TAU therapists ($t(24) = 2.8$, $p < .05$, $d = 1.13$), and TAU therapists reported a large effect size for proficiency in CBT compared to RFT therapists ($t(24) = -2.1$, $p = .05$, $d = .86$). No differences were found for self-rated proficiency in DC or MI.

As described above, four multilevel mixed-effects regression models were conducted to examine the main effects of proficiency, and its interaction with study condition, on the three ITT-ABP scales, controlling for covariates. A trend-level condition by proficiency interaction was found for DC proficiency predicting Drug Counseling Interventions ($B = -.40$, $SE = .21$, $z = -1.89$, $p = .06$, $d = .69$). Random effects for this model were significant for therapist (coefficient = $.49$, $SE = .08$) and client (coefficient = $.15$, $SE = .02$). This interaction was probed by splitting the sample on condition and re-conducting the regression to examine the within-condition effects of DC proficiency on Drug Counseling Interventions. Probing the interaction revealed that therapist proficiency in DC predicted greater use of Drug Counseling Interventions in the TAU condition only ($B = .42$, $SE = .14$, $z = 3.03$, $p < .01$, $d = .50$). Random effects for this model were significant for therapist (coefficient = $.61$, $SE = .12$) and client (coefficient = $.13$, $SE = .03$). A similar interaction was found for CBT proficiency predicting MI/CBT Interventions ($B = -.65$, $SE = .39$, $z = -1.65$, $p < .10$, $d = .79$), with significant random effects for therapist (coefficient = $.67$, $SE = .11$) and client (coefficient = $.24$, $SE = .04$ for client). Probing this interaction revealed that, similar to DC, therapist proficiency in CBT predicted greater use of MI/CBT Interventions in the TAU condition only ($B = .58$, $SE = .28$, $z = 2.06$, $p < .05$, $d = .61$), with significant random effects for therapist (coefficient = $.69$, $SE = .17$) and client (coefficient

= .35, SE = .09). No effects were found for FT proficiency predicting use of Family Interventions or for MI proficiency predicting MI/CBT Interventions.

Discussion

Study results provide strong support for factor and discriminant validity and preliminary support for convergent validity of the ITT-ABP, a therapist self-report measure of treatment techniques associated with four evidence-based behavioral approaches for adolescent conduct and substance use problems. Exploratory factor analysis of the ITT-ABP identified three clinically coherent scales: Family Interventions, MI/CBT Interventions, and Drug Counseling Interventions. These derived scales generally correspond with the four observational fidelity scales (FT, MI, CBT, DC) that were original sources for the study measure, albeit with a combined MI/CBT factor (discussed below), and the structural validity of all three scales was substantiated by confirmatory factor analysis. All scales demonstrated robust internal consistency across the whole sample, within treatment condition, and within different types of clinical sites in the TAU condition (in cases where the clinic type matched the EBP type), attesting to the within-factor reliability of the core items constituting each scale. As predicted, therapists working at the family-oriented RFT site reported significantly higher scores on the Family Interventions scale than therapists in the TAU condition; this difference remained large even when RFT was compared only to the two CMHC-type sites contained in TAU. Contrary to expectations, there were no between-condition differences in either MI/CBT scores or Drug Counseling scores.

There was only partial confirmation of the convergent validity of the ITT-ABP with a therapist-report measure of proficiency (theoretical allegiance and perceived skill) in each of the four approaches. In the TAU condition only, higher baseline proficiency in the DC approach predicted greater utilization of Drug Counseling Interventions when treating study cases, and higher CBT proficiency predicted greater MI/CBT scores. However, FT proficiency did not predict utilization of Family Interventions, nor did MI proficiency predict MI/CBT scores. Note that self-reported proficiency scores for FT and MI failed to predict their corresponding ITT-ABP scale scores despite benefitting from the psychometric leverage of common source bias and self-fulfilling prophecy bias. The disconnect between therapist views of their allegiance and skills in FT and MI, versus their self-reported implementation of FT and MI techniques when treating study cases, could be the result of measurement shortcomings of the self-reported proficiency variable, poor reliability among therapists in judging their own use of FT and MI interventions (discussed below), a real-world discrepancy between therapist proclivity to follow these broad approaches versus their decisions to apply their specific techniques when treating diverse clients in routine care, or some combination.

The EBP data reported by study clinicians working in usual care generally conformed to expectations set by controlled research on these four approaches. As a prime example, MI did not emerge as an autonomous treatment approach but rather as part of a treatment package with CBT. Packaging MI and CBT is a common strategy in behavioral treatment studies with adolescents (e.g., Battjes et al. 2004; Dennis et al. 2004), though MI has also proven effective as a stand-alone brief intervention for teens in a variety of settings (e.g., Breslin et al. 2002; D'Amico et al. 2008). Because study analyses collapsed ITT-ABP scores across sessions, it was not possible to determine whether MI techniques were implemented more densely in earlier sessions and CBT more densely in later sessions, a sequence prescribed in studies of multicomponent MI/CBT models. One hallmark CBT technique, Homework assignment, did not load on the MI/CBT factor or any other factor, similar to a previous finding that homework was rarely assigned during routine behavioral services for children with disruptive behavior disorders (Garland et al. 2010a, b).

Also as expected, therapists working in the family-oriented RFT clinic reported higher proficiency in the FT approach and favored family interventions over the other EBPs, though they reported using a sizable dose of MI/CBT as well. Conversely, therapists in non-family-oriented CMHCs and psychiatry clinics clearly favored MI/CBT but relied to some extent on family interventions as well. Perhaps most intriguing was the entirely eclectic mix reported by therapists in the addictions clinic. Although data were collected from only 3 therapists for 44 sessions at one addictions site, the high scores tallied at that site for all three scales heighten interest in the prevailing treatment curricula and therapeutic processes characterizing group-based treatments for adolescent substance use. By the same token, the fact that the mean scores on all three ITT-ABP scales at that site surpassed all scores from every other site (see Table 2) raises questions about the reporting accuracy of site therapists—or perhaps the reliability of the Drug Counseling items—that can be answered only via comparison to observational scores.

It is difficult to determine whether reported mean scores for the ITT-ABP scales represent a moderate or a low dose of the respective EBPs. In one of the very few observational investigations of usual care, three affiliated studies of community therapists treating childhood disruptive behaviors (Brookman-Frazee et al. 2010; Garland et al. 2010a, b; Hurlburt et al. 2010) converge in declaring that the practitioner sample evidenced a wide breadth but shallow depth (i.e., low intensity) of bona fide treatment techniques. In the current study, the average score for Family Interventions in RFT, and also for MI/CBT Interventions in TAU, was approximately 2.5. This mark falls between the anchors “A little bit” and “Moderately” on the ITT-ABP thoroughness/frequency scale, which is perhaps fairly descriptive. However, because scale scores were averaged across multiple items, it is entirely possible that in any given session, a mean score of 2.5 included one or more individual items that received a high-end score. There was also a considerable degree of variability in ITT scale scores both within therapist and within client, as expected in a usual care sample. To render an informed decision about EBP dosage in this sample, future studies will need to examine individual item scores and, as importantly, benchmark ITT-ABP scores against implementation data yielded by the source EST fidelity measures during controlled trials that carefully titrated intervention dosage.

Study Strengths and Limitations

This study examined the construct validity of a new measure of four specific EBPs. Focus groups with participating therapists indicated that the EBPs were prominently featured in their everyday practices. Still, the ITT-ABP was not designed to capture a broad and representative array of credible usual-care interventions for the sampled population. That is, the study did not attempt to open the black box of usual care but instead to shed light on a selected set of interventions allegedly within the box.

There were several study strengths. The study sampled community therapists operating in unadulterated clinical settings, that is, without benefit of extramural resources of any kind. In addition, adolescent clients were diagnosed with an array of clinical disorders, and comorbidity was the norm. These are conditions of high ecological validity that affirm the generalizability of study findings to real-world practice. Instrument development for the ITT-ABP was built upon the strongest possible foundation: empirically validated observational measures of empirically supported treatments that are widely endorsed for the target sample. The percentage of sessions for which study therapists in each condition submitted corresponding ITT-ABP checklists was promising (71 %) and likely underestimates what might be achieved in settings wherein QA reporting is mandatory (rather than voluntary, as in the current study). The study employed rigorous data analysis procedures, including statistical controls for nested data and for client and study design characteristics that would otherwise confound data interpretation. This includes multilevel

mixed-effects analyses that statistically controlled for nesting effects related to imbalance in number of checklists submitted per therapist and permitted inclusion of all available data points (checklists) to adequately power splitsample factor analyses.

There were also several limitations. The number of participating sites ($n = 6$) was too small to allow for generalizable claims about the practice habits of UC providers on a national scale. Indeed, the study did not contain a large enough sample of clinics to control or test for site differences, which are surely meaningful for understanding EBP implementation. The convergent validity analyses are considered preliminary only, due to the small number of therapists ($n = 26$) who provided treatment proficiency (i.e., allegiance and skill) data. Analyses controlled for therapist effects but did not investigate therapist differences in ITT-ABP scores due to lack of power. Also, therapists in RFT submitted a greater number of ITT-ABP checklists per client ($M = 13.2$) than did TAU ($M = 8.1$), which may be due to the larger critical mass of having 15 therapists located in one site versus 17 therapists spread across five sites; or instead, considering the greater submission variability in RFT ($SD = 13.3$) versus TAU ($SD = 8.1$), may be due more to a handful of prolific RFT data submitters. Because analyses controlled for therapist effects, these submission discrepancies do not endanger study results or interpretation. Future ITT-ABP studies should address therapist effects, as well as client and dose effects, on EBP implementation.

Certainly a major limitation of the study is the absence of observational data to verify the reliability of therapist self-reports. Unfortunately, studies attempting to confirm therapist reports of their own EBP implementation via observational ratings have produced disappointing results. Studies with adult populations have logged modest to weak correspondence between nonparticipant raters and therapist reports of fidelity to motivational interviewing (Martino et al. 2009; Miller et al. 2004) and cognitive-behavioral interventions (Brosan et al. 2008; Carroll et al. 1998). In the youth treatment arena, Hurlburt et al. (2010) found that observational coders reported substantially less occurrence and lower intensity of EBPs compared to therapist report in front-line mental health care. The ultimate validity of the ITT-ABP depends largely on establishing concurrence between (subjective) therapist reports and (objective) rater reports of implementation, to be pursued in subsequent studies.

A final step in ITT-ABP measure development will be testing predictive validity for targeted client outcomes. There is little knowledge about whether EBPs directly influence outcome in front-line settings, and virtually no knowledge about which implementation processes are, and are not, essential for producing key effects. However, even model-specific EST techniques, which have received the most attention to date, do not routinely predict outcome (Perepletchikova and Kazdin 2005), and when they do, the effect sizes are typically small (Webb et al. 2010) and the relation may be curvilinear in nature (e.g., Hogue et al. 2008a, b). It remains to be seen whether significant statistical correlations between EBP implementation data and client outcome data will be a required feature, or a desired but difficult-to-reach goal, of QA instrument development. That said, there can be no doubt that reliable QA tools are needed to ensure EBP fidelity, and that EBP fidelity is a necessary component of quality behavioral health services.

The ITT-ABP is limited to items representing specific treatment techniques, which are arguably the “active ingredients” of behavioral interventions. Yet there remains a dearth of research on the “contours” of implementation (Schoenwald et al. 2011) defined by the parameters of a given treatment (i.e., service delivery aspects of implementation: to whom, where, how often) and by prescribed treatment themes and session content (Garland et al. 2010a, b; Hogue et al. 2004). To make headway in developing efficient EBP measures for real-world application, it is critical to explore the feasibility of various methods for assessing

multiple dimensions of implementation (Schoenwald 2011). Asking therapists to judge the (more) readily defined targets and foci of their interventions, rather than discrete techniques that are often multifaceted and interwoven, sets the measurement bar a notch lower, which might engender improved reliability. Along these lines Kelley et al. (2010) developed a brief therapist-report measure of session focus that showed acceptable internal consistency and distinguished between clinician versus client influences on session content in usual care, though it has not been validated with observational data. Also, because validated observational measures are available for a limited number of EBPs at this time, the ongoing development of therapist-report QA instruments will need to rely on methods other than those of the current study.

Study Implications

It will not be possible to transport evidence-based treatments from the lab to the field until QA tools are available to support and monitor EBP fidelity in a cost-efficient and clinically relevant manner. The therapist-report ITT-ABP holds great promise to address this need for adolescent behavior problems—oppositionality, conduct disorder, substance misuse, and the like—for which cognitive-behavioral therapy, family therapy, motivational interviewing, and drug counseling are cornerstone approaches that are also widely endorsed by practicing clinicians (Cook et al. 2010; Gifford et al. 2012). Study findings indicate that the ITT-ABP has strong construct validity and clinical coherence, which is a necessary first step toward validation though still short of the finish line: verification that providers can reliably report on their own utilization of EBPs for ABP by means of post-session checklists. If this proves to be the case, it could advance the technology of QA tool development by illustrating that self-report fidelity checklists derived from validated measures of ESTs can be successfully retrofitted for usual care. This would open the door for developing a myriad of QA measures to support EBP implementation for almost every variety of clinical population.

Acknowledgments

Preparation of this article was supported by grant R01DA023945 from the National Institute on Drug Abuse. The authors would like to acknowledge the dedicated work of the clinical research staff for this project: Molly Bobek, Candace Johnson, and Emily Lichvar. We are also extremely grateful to the community therapists working at six study sites who generously agreed to participate in a research study, attend two focus groups and one instrument training, and complete numerous post-session checklists in the hopes of advancing the clinical science in their field.

References

- Aarons G, Hurlburt M, Horowitz S. Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Administration and Policy in Mental Health and Mental Health Services Research*. 2011; 38:4–23. [PubMed: 21197565]
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 4th edn.. APA; Washington, DC: 2000.
- Asparouhov T. Sampling weights in latent variable modeling. *Structural Equation Modeling*. 2005; 12(3):411–434.
- Ball SA, Martino S, Nich C, Frankforter TL, Van Horn D, Crits-Christoph P, Woody GE, Obert JL, Farentinos C, Carroll KM. Site matters: Multisite randomized trial of motivational enhancement therapy in community drug abuse clinics. *Journal of Consulting and Clinical Psychology*. 2007; 75:556–567. [PubMed: 17663610]
- Battjes RJ, Gordon MS, O'Grady KE, Kinlock TW, Katz EC, Sears EA. Evaluation of a group-based substance abuse treatment program for adolescents. *Journal of Substance Abuse Treatment*. 2004; 27:123–134. [PubMed: 15450645]
- Bearsley-Smith C, Sellick K, Chesters J, Francis K. Treatment content in child and adolescent mental health services: Development of the treatment recording sheet. *Administration and Policy in Mental Health and Mental Health Services Research*. 2008; 35:423–435. [PubMed: 18679789]

- Becker S, Curry J. Outpatient interventions for adolescent substance abuse: A quality of evidence review. *Journal of Consulting and Clinical Psychology*. 2008; 76:531–543. [PubMed: 18665683]
- Bond GR, Becker DR, Drake RE. Measurement of fidelity of implementation of evidence based practices: Case example of the IPS fidelity scale. *Clinical Psychology: Science and Practice*. 2011; 18:126–141.
- Breslin C, Li S, Sdao-Jarvie K, Tupker E, Ittig-Deland V. Brief treatment for young substance users: A pilot study in an addiction treatment setting. *Psychology of Addictive Behaviors*. 2002; 16:10–16. [PubMed: 11934080]
- Brookman-Frazee L, Haine RA, Baker-Ericzen M, Zoffness R, Garland AF. Factors associated with use of evidence-based practice strategies in usual care youth psychotherapy. *Administration and Policy in Mental Health and Mental Health Services Research*. 2010; 37:254–269. [PubMed: 19795204]
- Brosan L, Reynolds S, Moore RG. Self-evaluation of cognitive therapy performance: Do therapists know how competent they are? *Behavioural and Cognitive Psychotherapy*. 2008; 36:581–587.
- Browne, M.; Cudeck, R. Alternative ways of assessing model fit. In: Bollen, K.; Long, J., editors. *Testing structural models*. Sage Publications; Newbury Park, CA: 1993.
- Carroll KM, Ball SA, Nich C, Martino S, Frankforter TL, Farentinos C, Kunkel LE, Mikulich-Gilbertson SK, Morgenstern J, Obert JL, Polcin D, Snead N, Woody GE. Motivational interviewing to improve treatment engagement and outcome in individuals seeking treatment for substance abuse: A multisite effectiveness study. *Drug and Alcohol Dependence*. 2006; 81:301–312. [PubMed: 16169159]
- Carroll KM, Nich C, Rounsaville BJ. Utility of therapist session checklists to monitor delivery of coping skills treatment for cocaine abusers. *Psychotherapy Research*. 1998; 8:307–320.
- Carroll KM, Nich C, Sifry R, Nuro KF, Frankforter TL, Ball SA, Rounsaville BJ. A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. *Drug and Alcohol Dependence*. 2000; 57:225–238. [PubMed: 10661673]
- Chambers D, Ringeisen H, Hickman E. Federal, state, and foundation initiatives around evidence-based practices for child and adolescent mental health. *Child and Adolescent Mental Health*. 2005; 14:307–327.
- Chorpita B, Daleiden E, Ebesutani C, Young J, Becker K, Nakamura B, Phillips L. Evidence-based treatments for children and adolescents: An updated review of indicators of efficacy and effectiveness. *Clinical Psychology: Science and Practice*. 2011; 18:154–172.
- Cohen, J. *Statistical power analysis for the behavioral sciences*. 2nd ed.. Erlbaum; Hillsdale, NJ: 1988.
- Cook JM, Biyanova T, Elhai J, Schnurr PP, Coyne JC. What do psychotherapists really do in practice? An internet study of over 2,000 practitioners. *Psychotherapy: Theory, Research, Practice, Training*. 2010; 47:260–267.
- D'Amico EJ, Miles JNV, Stern SA, Meredith LS. Brief motivational interviewing for teens at risk of substance use consequences: A randomized pilot study in a primary care clinic. *Journal of Substance Abuse Treatment*. 2008; 35:53–61. [PubMed: 18037603]
- Damschroder L, Hagedorn H. A guiding framework and approach for implementation research in substance use disorders treatment. *Psychology of Addictive Behaviors*. 2011; 25:194–205. [PubMed: 21443291]
- Dennis ML, Godley SH, Diamond G, Tims FM, Babor T, Donaldson J, et al. The Cannabis Youth Treatment (CYT) study: Main findings from two randomized trials. *Journal of Substance Abuse Treatment*. 2004; 27:197–213. [PubMed: 15501373]
- Diamond GM, Diamond GS, Hogue A. Attachment-based family therapy: Adherence and differentiation. *Journal of Marital and Family Therapy*. 2007; 33:177–191. [PubMed: 17437458]
- Eyberg SM, Nelson MM, Boggs SR. Evidence-based psychosocial treatments for children and adolescents with disruptive behavior. *Journal of Clinical Child and Adolescent Psychology*. 2008; 37:215–237. [PubMed: 18444059]
- Fixsen, DL.; Naom, SF.; Blasé, KA.; Friedman, RM.; Wallace, F. *Implementation research: A synthesis of the literature*. University of South Florida, Louis de la parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231); Tampa, FL: 2005.

- Garland AF, Bickman L, Chorpita BF. Change what? Identifying quality improvement targets by investigating usual mental health care. *Administration and Policy in Mental Health and Mental Health Services Research*. 2010a; 37:15–26. [PubMed: 20177769]
- Garland AF, Hurlburt MS, Brookman-Frazee L, Taylor RM, Accurso EC. Methodological challenges of characterizing usual care psychotherapeutic practice. *Administration and Policy in Mental Health and Mental Health Services Research*. 2010b; 37:208–220. [PubMed: 19757021]
- Gifford EV, Tavakoli S, Weingardt KR, Finney JW, Pierson HM, Rosen CS, Curran GM. How do components of evidence-based psychological treatment cluster in practice?: A survey and cluster analysis. *Journal of Substance Abuse Treatment*. 2012; 42(1):45–55. [PubMed: 21943809]
- Hawkins EH. A tale of two systems: Co-occurring mental health and substance abuse disorders treatment for adolescents. *Annual Review of Psychology*. 2009; 60:197–227.
- Hedeker D, Gibbons RD, Flay BR. Random-effects regression models for clustered data with an example from smoking prevention research. *Journal of Consulting and Clinical Psychology*. 1994; 62:757–765. [PubMed: 7962879]
- Henggeler SW, Sheidow AJ, Cunningham PB, Donohue BC, Ford JD. Promoting the implementation of an evidence-based intervention for adolescent marijuana abuse in community settings: Testing the use of intensive quality assurance. *Journal of Clinical Child and Adolescent Psychology*. 2008; 37:682–689. [PubMed: 18645758]
- Hepner KA, Azocar F, Greenwood GL, Miranda J, Burnam MA. Development of a clinician report measure to assess psychotherapy for depression in usual care settings. *Administration and Policy in Mental Health and Mental Health Services Research*. 2010; 37:221–229. [PubMed: 20186569]
- Hogue A. When technology fails: Getting back to nature. *Clinical Psychology: Science and Practice*. 2010; 17:77–81. [PubMed: 20396629]
- Hogue A, Dauber S, Chinchilla P, Fried A, Henderson CE, Inclan J, Liddle HA. Assessing fidelity in individual and family therapy for adolescent substance abuse. *Journal of Substance Abuse Treatment*. 2008a; 35:137–147. [PubMed: 17997268]
- Hogue A, Dauber S, Samuolis J, Liddle H. Treatment techniques and outcomes in multidimensional family therapy for adolescent behavior problems. *Journal of Family Psychology*. 2006; 20:535–543. [PubMed: 17176187]
- Hogue A, Henderson CE, Dauber S, Barajas P, Fried A, Liddle H. Treatment adherence, competence, and outcome in individual and family therapy for adolescent behavior problems. *Journal of Consulting and Clinical Psychology*. 2008b; 76:544–555. [PubMed: 18665684]
- Hogue A, Liddle H. Family-based treatment for adolescent substance abuse: Controlled trials and new horizons in services research. *Journal of Family Therapy*. 2009; 31:126–154. [PubMed: 21113237]
- Hogue A, Liddle H, Dauber S, Samuolis J. Linking session focus to treatment outcome in evidence-based treatments for adolescent substance abuse. *Psychotherapy: Theory, Research, Practice, and Training*. 2004; 41:83–96.
- Hogue A, Liddle H, Rowe C, Turner R, Dakof GA, LaPann K. Treatment adherence and differentiation in individual versus family therapy for adolescent drug abuse. *Journal of Counseling Psychology*. 1998; 45:104–114.
- Hurlburt MS, Garland AF, Nguyen K, Brookman-Frazee L. Child and family therapy process: Concordance of therapist and observational perspectives. *Administration and Policy in Mental Health and Mental Health Services Research*. 2010; 37:230–244. [PubMed: 19902347]
- Kazak A, Hoagwood K, Weisz JR, Hood K, Kratochwill T, Vargas L, et al. A meta-systems approach for evidence-based practice for children and adolescents. *American Psychologist*. 2010; 65:85–97. [PubMed: 20141264]
- Kelley SD, de Andrade ARV, Sheffer E, Bickman L. Exploring the black box: Measuring youth treatment process and progress in usual care. *Administration and Policy in Mental Health and Mental Health Services Research*. 2010; 37:287–300. [PubMed: 20238155]
- Kolko DJ. Commentary: Studying usual care in child and adolescent therapy: It's anything but routine. *Clinical Psychology: Science and Practice*. 2006; 13:47–52.
- Martino S, Ball S, Nich C, Frankforter TL, Carroll KM. Correspondence of motivational enhancement treatment integrity ratings among therapists, supervisors, and observers. *Psychotherapy Research*. 2009; 19:181–193. [PubMed: 19396649]

- McDonald R, Ho M. Principles and practice in reporting structural equation analyses. *Psychological Methods*. 2002; 7:64–82. [PubMed: 11928891]
- McHugh RK, Barlow DH. The dissemination and implementation of evidence-based psychological treatments. *American Psychologist*. 2010; 65:73–84. [PubMed: 20141263]
- McHugh R, Murray H, Barlow DH. Balancing fidelity and adaptation in the dissemination of empirically supported treatments: The promise of transdiagnostic interventions. *Behavior Research and Therapy*. 2009; 47:946–953.
- Miller WR, Rose GS. Toward a theory of motivational interviewing. *American Psychologist*. 2009; 64:527–537. [PubMed: 19739882]
- Miller WR, Yahne CE, Moyers TB, Martinez J, Pirritano M. A randomized trial of methods to help clinicians learn motivational interviewing. *Journal of Consulting and Clinical Psychology*. 2004; 72:1050–1062. [PubMed: 15612851]
- Miranda J, Azocar F, Burnam MA. Assessment of evidence-based psychotherapy practices in usual care: Challenges, promising approaches, and future directions. *Administration and Policy in Mental Health and Mental Health Services Research*. 2010; 37:205–207. [PubMed: 20354778]
- Nakamura BJ, Daleiden EL, Mueller CW. Validity of treatment target progress ratings as indicators of youth improvement. *Journal of Child and Family Studies*. 2007; 16:729–741.
- Perepletchikova F. On the topic of treatment integrity. *Clinical Psychology: Science and Practice*. 2011; 18:148–153. [PubMed: 21769167]
- Perepletchikova F, Kazdin AE. Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice*. 2005; 12:365–383.
- Rabe-Hesketh, S.; Skrondal, A. *Multilevel and longitudinal modeling using Stata*. StataCorp LP; College Station, TX: 2005.
- Raudenbush, SW.; Bryk, AS. *Hierarchical linear models: Applications and data analysis methods*. 2nd ed.. Sage; Thousand Oaks, CA: 2002.
- Schoenwald SK. It's a bird, it's a plane, it's fidelity measurement in the real world. *Clinical Psychology: Science and Practice*. 2011; 18:142–147. [PubMed: 21691439]
- Schoenwald SK, Garland AF, Chapman JE, Frazier SL, Sheidow AJ, Southam-Gerow MA. Toward the effective and efficient measurement of implementation fidelity. *Administration and Policy in Mental Health and Mental Health Services Research*. 2011; 38:32–43. [PubMed: 20957425]
- Schoenwald SK, Sheidow AJ, Chapman JE. Clinical supervision in treatment transport: Effects on adherence and outcomes. *Journal of Consulting and Clinical Psychology*. 2009; 77:410–421. [PubMed: 19485583]
- Waldron HB, Turner CW. Evidence-based psychosocial treatments for adolescent substance abuse. *Journal of Clinical Child and Adolescent Psychology*. 2008; 37:238–261. [PubMed: 18444060]
- Webb CA, DeRubeis RJ, Barber JP. Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology*. 2010; 78:200–211. [PubMed: 20350031]
- Weersing RV, Weisz JR, Donenberg GR. Development of the therapy procedures checklist: A therapist-report measure of technique use in child and adolescent treatment. *Journal of Clinical Child Psychology*. 2002; 31:168–180.
- Winters KC, Stinchfield R, Latimer WW, Lee S. Long-term outcome of substance-dependent youth following 12-step treatment. *Journal of Substance Abuse Treatment*. 2007; 33:61–69. [PubMed: 17588490]
- Winters KC, Stinchfield RD, Opland E, Weller C, Latimer WW. The effectiveness of the Minnesota Model approach in the treatment of adolescent drug abusers. *Addiction*. 2000; 95:601–612. [PubMed: 10829335]
- Zazzali JL, Sherbourne C, Hoagwood K, Greene D, Bigley MF, Sexton T. The adoption and implementation of an evidence based practice in child and family mental health services organizations: A pilot study of Functional Family Therapy in New York State. *Administration and Policy in Mental Health and Mental Health Services Research*. 2008; 35:38–49. [PubMed: 17985232]

Table 1

Exploratory factor analysis of ITT-ABP: three factor solution

	Theoretical Factor*	Factor 1	Factor 2	Factor 3
Factor Eigenvalue		9.809	4.060	2.745
Percent Variance Accounted for		36.3 %	15.04 %	10.2 %
Factor 1: Drug Counseling Interventions				
Explored positive and negative effects of substance use and the pros and cons of abstinence	DC	.974	-.174	-.030
Discussed cravings, triggers, and high-risk situations that lead to current or future use	DC	.971	-.264	.128
Discussed the disease concept of addiction	DC	.953	.027	.015
Advocated a goal of abstinence from drug use	DC	.929	.007	.088
Explored or confronted the client's denial in relation to drug use or consequences of use	DC	.893	.096	-.016
Encouraged client to participate in 12-step meetings and activities	DC	.873	.303	-.111
Emphasized spirituality as an important component of recovery	DC	.791	.116	-.084
Discussed one or more of the 12 steps	DC	.781	.204	-.044
Helped client develop friends, relationships, and social activities that are non-drug related	CBT	.444	.069	.002
Assigned homework and/or reviewed homework from previous session**	CBT	.379	.063	.163
Factor 2: MI/CBT Interventions				
Facilitated the client's awareness of discrepancies between current problematic behaviors and future goals	MI	-.034	.806	.082
Explored client concerns about problematic behavior, readiness to change behavior, and optimism about success	MI	-.063	.799	.027
Worked on client's commitment to a plan for changing problematic behavior, including discussion of impediments to change	MI	-.017	.793	.129
Affirmed the client's ability to change problematic behavior and praised change efforts	MI	.053	.705	.170
Used cognitive monitoring techniques to call attention to client self-talk and how thoughts can cause feelings and behavior	CBT	.181	.696	-.146
Taught client new problem-solving, coping, or communication skills	CBT	.294	.546	-.074
Repeated the client's words, paraphrased client statements, or made reflective summary statements back to the client	MI	-.054	.526	.104
Utilized behavioral interventions (e.g., formalized treatment planning, reward systems), or taught relaxation exercises (e.g., meditation)	CBT	.275	.523	-.098
Factor 3: Family Interventions				
Worked on enhancing communication and attachment between family members	FT	.047	.204	.734
Targeted intervention efforts at a family member participating in the session	FT	-.157	-.023	.702
Arranged, coached, and helped process an in-session family interaction	FT	-.101	-.269	.689
Discussed core relational family themes that underlie everyday events (e.g., love, trust, respect, independence)	FT	.123	.252	.635
Discussed parental monitoring and family rules/caretaking with the adolescent and/or caregiver	FT	.072	.014	.617
Established definite theme or agenda at beginning of session	CBT	.092	.285	.473
Shared information about normative adolescent development	FT	.064	.279	.464
Worked individually with adolescent or caregiver to prepare for family session (same day or future session)	FT	.277	-.117	.456
Emphasized equality and collaboration in the therapeutic relationship versus "therapist in	MI	.222	.289	.364

	Theoretical Factor*	Factor 1	Factor 2	Factor 3
charge ^{**,*}				

* *DC* Drug Counseling, *MI* Motivational Interviewing, *CBT* Cognitive-Behavioral Therapy, *FT* Family Therapy

** These items did not load above .40 on any factor and were dropped from further analysis

Table 2

Condition differences in average scores and internal consistency of ITT-ABP subscales

	<i>N</i>	Mean (SD)	Alpha
RFT (1 clinic)			
Drug Counseling Interventions	505	1.18 (.28)	.56
MI/CBT Interventions	505	2.31 (.72)	.83
Family Interventions	505	2.68 (.70)	.71
TAU (5 clinics)			
Drug Counseling Interventions	317	1.40 (.84)	.95
MI/CBT Interventions	317	2.45 (.95)	.90
Family Interventions	317	2.04 (.72)	.80
TAU: CMHCs (2 clinics)			
Drug Counseling Interventions	195	1.03 (.09)	.03
MI/CBT Interventions	195	2.22 (.92)	.90
Family Interventions	195	1.92 (.68)	.79
TAU: Addictions Clinic (1 clinic)			
Drug Counseling Interventions	44	3.42 (.50)	.57
MI/CBT Interventions	44	3.61 (.56)	.69
Family Interventions	44	2.80 (.64)	.73
TAU: Outpatient Psychiatry (2 clinics)			
Drug Counseling Interventions	78	1.19 (.17)	.11
MI/CBT Interventions	78	2.40 (.73)	.86
Family Interventions	78	1.90 (.61)	.76

RFT Routine Family Therapy, *TAU* Treatment as Usual, *MI* Motivational Interviewing, *CBT* Cognitive-Behavioral Therapy

Table 3

Condition differences in therapist self-reported proficiency

	RFT (n = 10) Mean (SD)	TAU (n = 16) Mean (SD)	t	p	d
FT Proficiency	3.7 (.88)	2.7 (.80)	2.8	.01	1.13
MI Proficiency	2.5 (1.0)	2.5 (1.0)	-.08	.94	.03
CBT Proficiency	2.6 (.77)	3.2 (.71)	-2.1	.05	.86
DC Proficiency	2.0 (.94)	2.1 (1.2)	-.14	.89	.06

RFT Routine Family Therapy, TAU Treatment as Usual, FT Family Therapy, MI Motivational Interviewing, CBT Cognitive-Behavioral Therapy, DC Drug Counseling