Journal of Consulting and Clinical Psychology
2008, Vol. 76, No. 4, 544–555

Copyright 2008 by the American Psychological Association
0022-006X/08/$12.00 DOI: 10.1037/0022-006X.76.4.544

Treatment Adherence, Competence, and Outcome in Individual and Family Therapy for Adolescent Behavior Problems

Aaron Hogue
National Center on Addiction and Substance Abuse at Columbia University

Craig E. Henderson
Sam Houston State University

Sarah Dauber, Priscilla C. Barajas, and Adam Fried
National Center on Addiction and Substance Abuse at Columbia University

Howard A. Liddle
University of Miami Miller School of Medicine

This study examined the impact of treatment adherence and therapist competence on treatment outcome in a controlled trial of individual cognitive–behavioral therapy (CBT) and multidimensional family therapy (MDFT) for adolescent substance use and related behavior problems. Participants included 136 adolescents (62 CBT, 74 MDFT) assessed at intake, discharge, and 6-month follow-up. Observational ratings of adherence and competence were collected on early and later phases of treatment (192 CBT sessions, 245 MDFT sessions) by using a contextual measure of treatment fidelity. Adherence and competence effects were tested after controlling for therapeutic alliance. In CBT only, stronger adherence predicted greater declines in drug use (linear effect). In CBT and MDFT, (a) stronger adherence predicted greater reductions in externalizing behaviors (linear effect) and (b) intermediate levels of adherence predicted the largest declines in internalizing behaviors, with high and low adherence predicting smaller improvements (curvilinear effect). Therapist competence did not predict outcome and did not moderate adherence–outcome relations; however, competence findings are tentative due to relatively low interrater reliability for the competence ratings. Clinical and research implications for attending to both linear and curvilinear adherence effects in manualized treatments for behavior disorders are discussed.

Keywords: treatment adherence, therapist competence, adolescent substance use, cognitive–behavioral therapy, multidimensional family therapy

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Rigorous fidelity monitoring and evaluation are required elements of efficacy research on manual-based behavioral interventions (Carroll, Kadden, Donovan, Zweben, & Rounsaville, 1994), and fidelity research is rapidly becoming a centerpiece of treatment dissemination as well. Some evidence has indicated that strong fidelity to empirically based interventions may be essential for producing treatment effects in real world settings. For example, Henggeler and colleagues (Henggeler, Melton, Brondino, Scherer, & Hanley, 1997; Henggeler, Pickrel, & Brondino, 1999) found that fidelity to multisystemic therapy for delinquent adolescents was poor when community therapists implemented the model without ongoing supervision from model experts; moreover, poor fidelity was linked to worse outcomes compared with results from efficacy studies. Interestingly, Morgenstern, Morgan, McCrady, Keller, and Carroll (2001) showed that whereas community practitioners intensively trained in cognitive–behavioral therapy for adult substance abuse could reach fidelity and outcome benchmarks set by research therapists, control group practitioners with no additional training reached the same outcome benchmarks. As dissemination research continues to mature, it seems certain that fidelity issues will remain a priority for treatment developers, program administrators, and policymakers.

Research on the link between fidelity and outcome is therefore positioned to make a major contribution to treatment development and dissemination efforts. Two aspects of treatment fidelity have received the most attention to date (Waltz, Addis, Koerner, & Jacobson, 1993): *adherence* (or, integrity), which refers to the extensiveness or dosage of model-prescribed intervention techniques implemented in session; and *competence*, which refers to the quality or skill with which interventions are delivered. A sizable number of studies, most of them using observational coding methods, have examined the association between adherence and outcome in research settings, and the results are mixed (Miller

Aaron Hogue, Sarah Dauber, Priscilla C. Barajas, and Adam Fried, National Center on Addiction and Substance Abuse at Columbia University; Craig Henderson, Department of Psychology, Sam Houston State University; Howard A. Liddle, Department of Epidemiology and Public Health and Department of Psychology, Center for Treatment Research on Adolescent Drug Abuse, University of Miami Miller School of Medicine.

Preparation of this article was supported by Grants R01 DA14571 and P50 DA07697 from the National Institute on Drug Abuse. We gratefully acknowledge Jaime Inclan of the Roberto Clemente Family Guidance Center and Robert H. Reiner of Behavioral Associates in New York City for supporting the treatment fidelity coding groups, Leyla Faw and John J. Cecero for their research contributions, and Dustin L. Jones for his work on interpreting and graphing the curvilinear adherence–outcome relationship.

Correspondence concerning this article should be addressed to Aaron Hogue, Associate Director, Health and Treatment Research, National Center on Addiction and Substance Abuse at Columbia University, 633 Third Avenue, 19th floor, New York, NY 10017. E-mail: ahogue@casacolumbia.org

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

& Binder, 2002; Perepletchikova & Kazdin, 2005). Some have concluded that strong adherence reflects therapist rigidity and overreliance on technique, which undermines development of a strong therapeutic relationship (Castonguay, Goldfried, Wiser, Raue, & Hayes, 1996; Henry, Strupp, Butler, Scacht, & Binder, 1993). Others have found that greater adherence predicts better outcome (Frank, Kupfer, Wagner, McEachran, & Cornes, 1991; Huey, Henggeler, Brondino, & Pickrel, 2000) and that adherence in early treatment sessions either predicts (Feeley, DeRubeis, & Gelfand, 1999) or is predicted by (Barber, Crits-Christoph, & Luborsky, 1996) early symptom improvement.

Few studies have examined whether competence predicts outcome, and here too the findings are inconsistent, with competence showing moderate effects in some studies (e.g., Barber et al., 1996; Shaw et al., 1999) and no effects in others (e.g., Barber et al., 2006). The paucity of research in this area may reflect the difficulties inherent in mounting a rigorous and clinically valid assessment of therapist competence. To assess competence, one must judge the skill of the therapist in the given model, the appropriateness and timing of interventions, and the degree of responsiveness to client behaviors (Stiles, Honos-Webb, & Surko, 1998). Unbiased judges with this level of sophistication are hard to recruit. Moreover, competence assessment should be grounded in thorough knowledge of the client and therapeutic context. Ideally this entails viewing multiple sessions per case, which requires a significant expenditure of time and resources (Waltz et al., 1993).

Fidelity studies also face the methodological challenge of demonstrating that fidelity–outcomes relations are not confounded by third-variable influences. Client characteristics (e.g., symptom severity, motivation to change) or relationship factors (e.g., therapeutic alliance) may be related to both fidelity and outcome and thereby account indirectly for observed fidelity–outcome effects. Perepletchikova and Kazdin (2005) presented several strategies for improving tests of direct fidelity effects, including the following: investigate empirically supported treatments with known effectiveness, have non-participant judges rate randomly selected sessions by using validated fidelity measures, and assess related treatment processes to control third-variable influences. A recent exemplary study by Barber et al. (2006) tested both linear and curvilinear (i.e., quadratic) adherence effects in drug counseling for cocaine users. Results showed that intermediate adherence, representing a balance between protocol integrity and clinically flexible deviation, predicted greater improvement in drug use and depression symptoms than did high (rigid) adherence or low (lax) adherence. Unexpectedly, competence did not predict outcome directly, nor did it moderate adherence–outcome effects.

The current study examines fidelity–outcome relations in a controlled trial comparing individual cognitive–behavioral therapy (CBT) and multidimensional family therapy (MDFT) for adolescent substance use and related behavior problems (Liddle, Dakof, Turner, Henderson, & Greenbaum, in press). Two previous studies have explored treatment technique–outcome effects in this same clinical sample. In a small study ($N = 51$), Hogue, Liddle, Dauber, and Samuolis (2004) found that use of family-focused treatment techniques, but not adolescent-focused techniques, predicted reductions in marijuana use, externalizing symptoms, and internalizing symptoms across CBT and MDFT at treatment discharge. A follow-up study on the MDFT condition only (Hogue, Dauber, Samuolis, & Liddle, 2006) found that use of family techniques was related to decrease in internalizing and increase in family cohesion 1 year after treatment. Family techniques also predicted reduced externalizing and family conflict, but only when use of adolescent techniques was also high. Finally, use of adolescent techniques themselves predicted improved family outcomes at 1 year.

The current study investigated the impact of adherence and competence in CBT and MDFT on marijuana use, personal problems related to drug use, and internalizing and externalizing symptoms up to 6 months after treatment. These dependent variables were selected specifically because they had already demonstrated significant outcome effects in the clinical trial from which study cases were drawn (Liddle et al., in press). Following Barber et al. (2006), we hypothesized that adherence would show a curvilinear effect on outcome, with intermediate adherence levels predicting greater change than did either high or low levels. We hypothesized that competence would instead show a linear effect, with more competence predicting better outcome. We also explored the moderating effect of competence on adherence–outcome relations to determine whether adherence to protocol is more important for cases with relatively lower therapist competence.

This study advances treatment fidelity research on behavioral interventions in several ways. It is one of the first to examine the impact of therapist competence on treatment outcome in adolescents and to examine fidelity–outcome effects in a substance-abusing adolescent sample. We utilized a contextual method of fidelity assessment that featured observational ratings of molar treatment modules (rather than discrete treatment techniques) implemented across multiple sessions for each client. Potential third-variable influences on fidelity–outcome relations were addressed by controlling for level of therapeutic alliance. Previous research on this sample (Hogue, Dauber, Faw, Cecero, & Liddle, 2006) found no relation between early alliance and outcome in CBT. In MDFT, there were both direct effects—stronger parent alliance predicted declines in marijuana use and externalizing symptoms—and an "alliance shift" effect, whereby adolescents who improved from weaker early alliance to stronger mid-treatment alliance showed corresponding improvement in externalizing compared with those whose alliances deteriorated.

Method

The study was conducted under approval by the governing Internal Review Board. Active consent from caregivers and assent from adolescents were collected. Therapists provided active consent for their sessions to be judged for adherence and competence.

Participants

Clients. Clients included 136 substance using adolescents drawn from a randomized trial ($N = 224$) comparing CBT and MDFT (Liddle et al., in press). All youths in the trial were drug users, with 75% meeting *DSM–IV* criteria for cannabis dependence and 13% for cannabis abuse, 20% alcohol dependence and 4% alcohol abuse, and 13% other drug dependence and 2% other drug abuse. Clients could meet diagnostic criteria for more than one substance use disorder. Also, 79% met criteria for oppositional defiant and/or conduct disorder and 49% for a mood and/or anxiety disorder. Cases from this randomized trial were included in the current study (62 CBT, 74 MDFT) if they completed a baseline

and posttreatment assessment (discharge and/or 6-month follow-up) and at least one videotaped therapy session. The final study sample was 81% male adolescents with an average age of 15.5 years ($SD = 1.3$) and a range of 13–17 years. Ethnic composition was 70% African American, 20% European American, and 10% Hispanic American. Half were living in one-parent households, 14% with both biological parents, and 36% with various other arrangements. Yearly household income was less than $10,000 for 29% of families. Most adolescents were enrolled in school (76%) and on juvenile probation (63%) at intake, and 32% had been court ordered to receive treatment.

Sample bias analyses were conducted to determine whether the 136 participants selected for this study differed from the clinical trial sample of 224 on the demographic variables described above and intake values of all study outcomes. The only differences found were the study sample having higher scores on parent-report externalizing, $t(213) = -2.15$, $p < .05$; and youth-report externalizing, $t(215) = -2.07$, $p < .05$. Also, as expected the study sample attended more treatment sessions than did the trial sample, $t(220.9) = -9.50$, $p < .001$. For the study sample, cases completed an average of 12.3 sessions ($SD = 8.7$), with 59% of cases attending 8 or more. In the trial sample, cases completed an average of 8.7 sessions ($SD = 0.90$), and 20% never attended any treatment session.

Therapists. The nine therapists who delivered the treatments, four in CBT and five in MDFT, ranged in age from 29 to 54 years ($M = 40$). The CBT therapists (two women) included two African Americans and two European Americans. One had a master's degree and three had doctorates, with an average of 3.5 years ($SD = 1.7$) postgraduate experience in CBT. MDFT therapists (three women) included three African Americans and two European Americans. Four had a master's and one had a doctorate, with an average of 7.7 years ($SD = 4.5$) postgraduate experience in family therapy.

Treatments

Individual CBT. The CBT model for multiproblem adolescent substance users (Turner, 1992; Waldron & Kaminer, 2004) is based on a broadly defined cognitive–behavioral framework that emphasizes a harm-reduction approach to substance use. CBT has demonstrated efficacy for adult drug users in individual format (Crits-Christoph et al., 1999) and adolescent substance users in group format (Dennis et al., 2004) and individual format (Waldron, Slesnick, Brody, Turner, & Peterson, 2001). Initial sessions focus on prioritizing adolescent problems and constructing the treatment contract. The intensive cognitive–behavioral intervention program then focuses on increasing coping competence and reducing problematic behaviors. By using a modular approach, therapists select treatment strategies based on the needs of the individual adolescent, including health education, contingency contracting, self-monitoring, problem-solving skills, communication skills, identifying cognitive distortions, and increasing prosocial activities. Role rehearsal and homework assignments are utilized to practice and reinforce new skills. Final sessions focus on relapse prevention and maintenance of gains.

MDFT. MDFT (Liddle, 2002) is a developmental–ecological treatment for adolescent drug abuse that seeks to reduce symptoms and enhance developmental functioning by facilitating change in several behavioral domains. The model has proven efficacious with adolescent substance users in outpatient treatment (Dennis et al., 2004; Liddle et al., 2001, in press) and with early-stage adolescent users (Liddle, Rowe, Dakof, Ungaro, & Henderson, 2004). MDFT has four interdependent modules that target multiple aspects of adolescent and family functioning. The *adolescent module* aims to build a therapeutic alliance with the adolescent, improve problem-solving skills and social competence, and develop alternative behaviors to drug use. The *parent module* aims to build a therapeutic alliance with the parent, increase parental involvement with the adolescent, and improve parenting skills. The *interactional module* works with parents and adolescents conjointly to strengthen emotional attachments and patterns of communication. The *extrafamilial module* seeks to establish collaborative relationships among all social systems in which the adolescent participates (family, school, peer, recreational, juvenile justice).

Treatment fidelity procedures. Therapists were given study cases after 4 months of training and on achieving satisfactory levels of fidelity in pilot cases as judged by model developers. Therapists were supervised weekly by model experts via live individual supervision, videotape feedback, and group supervision. Both treatments prescribed office-based, weekly sessions conducted over 16–24 weeks. Treatment adherence to signature therapy techniques was previously established for both conditions in the original clinical trial by using a randomly selected subset of 36 cases total (Hogue et al., 1998).

Outcome Measures

Marijuana use frequency: Timeline follow-back. The timeline follow-back (Sobell & Sobell, 1996) measures quantity and frequency of daily consumption of substances by using a calendar and other memory aids to gather retrospective estimates. The timeline follow-back is reliable and valid for the measurement of alcohol consumption and cigarette and cannabis use (Breslin, Sobell, & Sobell, 1996). Criterion validity has been established by comparing self- and collateral reports as well as self-reports and records of verifiable events such as hospitalizations and jail stays (Fals-Stewart, O'Farrell, Freitas, McFarlin, & Rutigliano, 2000). This study examined the number of days in the past 30 during which the adolescent smoked marijuana, the primary drug of use in this sample. Across both conditions, mean number of marijuana use days was 11.9 ($SD = 12.5$) at baseline, 6.8 (11.9) at post, and 5.2 (8.6) at follow-up.

Drug use problems: Personal Experience Inventory (PEI). The PEI is a multiscale self-report measure assessing drug use problem severity and psychosocial risk (Winters, Latimer, & Stinchfield, 2002). This study used the total score of the Personal Involvement with Chemicals scale, a 29-item measure focusing on the psychological and behavioral involvement in substance use and related problems in the previous 30 days. Widely used in applied research settings, it has shown excellent reliability (Cronbach's $\alpha = .84–.97$) and validity (scales significantly related to diagnostic ratings) in adolescents from diverse ethnic backgrounds. Across both conditions, the mean problem severity score was 28.9 ($SD = 17.9$) at baseline, 23.6 (19.2) at post, and 19.5 (18.9) at follow-up.

Externalizing and internalizing symptoms: Child Behavior Checklist (CBCL) and Youth Self-Report (YSR). The Revised CBCL (Achenbach, 1991a) is a parent self-report measure that

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

assesses children's behavior problems and social competencies. It contains groupings of Externalizing symptoms (delinquent and aggressive) and Internalizing symptoms (withdrawn, anxious/depressed, somatic complaints). In previous studies, researchers have obtained 1-week test–retest reliability estimates of .93 as well as interparent reliability of .66 for Internalizing and .80 for Externalizing (Achenbach, 1991a). Content and criterion validity are supported by the ability of CBCL items to discriminate between matched referred and non-referred youths (Achenbach, 1991a). The YSR (Achenbach, 1991b) is a youth-report version of the CBCL with equivalent psychometric properties. The current study included summary scales of CBCL and YSR externalizing symptoms and CBCL internalizing symptoms; YSR internalizing symptoms were not examined because they did not significantly improve in the original clinical trial. Across both conditions, the mean score for CBCL externalizing was 25.5 (SD = 11.9) at baseline, 20.9 (12.7) at post, and 18.6 (12.2) at follow-up; for YSR externalizing, 19.1 (SD = 9.2) at baseline, 17.1 (8.5) at post, and 15.9 (8.9) at follow-up; and for CBCL internalizing, 11.5 (SD = 8.0) at baseline, 9.4 (7.9) at post, and 8.1 (8.0) at follow-up.

Process Measures

Therapist Behavior Rating Scale—Competence (TBRS–C). The TBRS–C is an observational measure of treatment adherence and therapist competence for individual CBT and MDFT for adolescent substance use and related behavior problems. Scale items represent the molar treatment modules of the given model, which are composed of multiple integrated intervention techniques that typically extend across several sessions (Diamond & Diamond, 2002). Items are scored on a 7-point Likert-type scale with the following anchors: 1 (*not at all*), 3 (*somewhat*), 5 (*considerably*), and 7 (*highly*). Each item receives a separate score for adherence and competence. Adherence ratings estimate the thoroughness and frequency with which interventions are executed. Competence ratings estimate the technical quality of interventions (skillfulness) and the timing and appropriateness of interventions for the given client and situation (responsiveness). Each TBRS–C item assists judges in making competence assessments for that module by describing treatment context considerations (e.g., client interpersonal style, treatment phase) and keys to competent module implementation.

The TBRS–C contains five molar treatment modules for individual CBT: Establishing a Working Relationship, Drug Use Monitoring and Harm Reduction (exemplary techniques: analysis of drug use behavior, refusal skills, and moderated use), Behavioral Skills Training (communication skills, decision making and problem solving, anger management, role playing, relaxation training), Cognitive Therapy Techniques (cognitive monitoring and change strategies, coping with drug use thoughts), and Increasing Prosocial Behavior. It also contains four molar treatment modules for MDFT (see Table 2): Adolescent Interventions (exemplary techniques: building and maintaining adolescent alliance, mapping ecological influences on prosocial and antisocial behavior, exploring drug use behaviors and consequences), Parent Interventions (building and maintaining parent alliance, reinforcing attachment and resuscitating hope, enhancing parental monitoring and discipline), Family Interaction Interventions (meeting individually with family members to prepare for family sessions, resolving parent–adolescent impasses, promoting positive family dialogue), and Extrafamilial Interventions (school and vocational interventions, juvenile justice interventions). For the current study, scale average scores were created for each condition by averaging the final scores for the treatment modules (five items for CBT, four for MDFT) for each session; the adherence and competence study variables were then created by averaging the scale average scores across all available sessions for each case.

The psychometric properties of the TBRS–C, including construct and discriminant validity, are presented in detail elsewhere (Hogue et al., in press). A summary of means, standard deviations, and interrater reliabilities for each TBRS–C item is contained in Table 1. According to Cicchetti's (1994) criteria for classifying the utility of ICC magnitudes, below .40 is poor, .40 to .59 is fair, .60 to .74 is good, and .75 to 1.00 is excellent. Both the CBT and MDFT items demonstrated good-to-excellent interrater reliability for adherence but only fair-to-poor reliability for competence as measured by the intraclass correlation coefficient (ICC[1,2]; Shrout & Fleiss, 1979). Reliability for the five CBT modules ranged from ICC = .56 to .83 for adherence and ICC = .01 to .63

Table 1
Descriptive Statistics and Intraclass Correlation Coefficients (ICCs) for CBT and MDFT

	Adherence				Competence			
Type of therapy	*N*	*M*	*SD*	ICC	*N*	*M*	*SD*	ICC
CBT								
1. Establishing a Working Relationship	192	2.96	1.85	.83	90	4.38	0.91	.48
2. Drug Use Monitoring and Harm Reduction	192	2.89	1.24	.81	138	3.74	0.95	.63
3. Behavioral Skills Training	192	2.09	1.28	.62	56	3.71	0.84	.25
4. Cognitive Therapy Techniques	192	1.56	1.00	.73	30	3.77	0.87	.01
5. Increasing Prosocial Behavior	192	2.06	1.13	.56	57	3.53	0.95	.39
MDFT								
1. Adolescent Interventions	245	4.68	1.85	.73	188	5.40	1.06	.48
2. Parent Interventions	245	4.74	1.95	.79	189	5.62	0.96	.48
3. Family Interaction Interventions	245	3.28	1.89	.66	109	5.22	1.04	.29
4. Extrafamilial Interventions	245	2.02	1.35	.64	62	4.85	0.97	.15

Note. CBT = cognitive–behavioral therapy; MDFT = multidimensional family therapy.

for competence. Reliability for the four MDFT modules ranged from ICC = .64 to .79 for adherence and ICC = .15 to .48 for competence. Note that ICCs for the scale average competence scores, which were the variables used in fidelity–outcome analyses, were fair: ICC = .56 for CBT and .55 for MDFT. In CBT, the mean scale average score was 2.29 (SD = 0.41) for adherence and 3.83 (0.74) for competence; in MDFT, the mean score was 3.66 (0.56) for adherence and 5.43 (0.78) for competence. Because two separate groups of coders were used to rate CBT and MDFT tapes, it is not possible to decide whether the higher mean scores for MDFT fidelity reflect a true superiority in treatment fidelity (condition effect), a proclivity in the MDFT judges themselves to give higher scores (judge effect), or an interaction between the two. All four fidelity variables were normally distributed. The correlation between scale average adherence and competence was $r(62) = .44$ in CBT and $r(74) = .32$ in MDFT.

Internal consistency reliability (Cronbach's α) was not calculated for the TBRS–C, for two reasons. First, the two TBRS–C scales are designed to capture multifaceted therapy modules contained within multicomponent, flexibly delivered intervention models. Thus TBRS–C items are theoretically independent and not intended to represent a correlated set of discrete interventions composing a single, unified construct—that is, more work in one module does not predict more work in other modules in any given session. On the contrary, extensive work in one module more or less precludes extensive work in other modules for any given session. Second, because model-specific clinical expertise is needed to code therapist competence in a valid manner (Waltz et al., 1993), only family therapists rated MDFT tapes and only CBT therapists rated CBT tapes. This within-condition coding design (Startup & Shapiro, 1993) attenuates potential correlations among fidelity items from a given scale because the items are not collectively set in contrast to items from a competing treatment being rated by the same group of coders.

Vanderbilt Therapeutic Alliance Scale—Revised (VTAS–R). The VTAS–R is a 22-item version of the original Vanderbilt Therapeutic Alliance Scale (VTAS; Hartley & Strupp, 1983) that defines the therapeutic alliance as a collaborative and task-oriented relationship determined by client behaviors and therapist–client relationship characteristics. Each item is rated on a Likert-type scale ranging from 0 (*not at all*) to 5 (*a great deal*). The current study used ratings of therapist–adolescent alliance; raters coded entire sessions in which the teen was present for at least 15 minutes. In a previous study of the same client pool, Hogue, Dauber, Faw, et al. (2006) found that VTAS–R ratings yielded a single underlying factor with excellent interrater reliability and internal consistency in both conditions: ICC(1,2) = .90 and Cronbach's α = .98 in CBT; and ICC(1,2) = .83 and α = .97 in MDFT. Ratings of therapist–adolescent alliance were modestly correlated with scale average ratings of adherence, $r(71) = .28$ in CBT, $r(73) = .19$ in MDFT; and competence, $r(67) = .13$ in CBT, $r(73) = .40$ in MDFT.

Sampling Procedures

TBRS–C. Videotaped sessions were selected from Phase 1 of treatment (every study case) and from Phase 2 (when available). Phase 1 contained the first 2 available sessions between 1 and 5, so that judges could evaluate client presenting problems and early treatment developments as a context for coding later sessions. Phase 2 contained a randomly selected set of 3 consecutive sessions (when available) starting at session 6. Identical sampling procedures were used for both conditions. However, fewer sessions from the CBT condition were included in this study due to its somewhat higher treatment dropout rate in the original clinical trial: 36% of cases randomized to CBT dropped from treatment prior to session 6, compared with 31% in MDFT. In CBT, 192 sessions were selected from 62 cases. Due to early treatment dropout, 36% of cases had Phase 1 tapes only. Across the 192 sessions, 54% were Phase 1 tapes, 29% were Phase 2 tapes that fell between sessions 6 and 12, and 17% were Phase 2 tapes between 13 and 25. For Phase 1 sets, 62% contained the first 2 sessions of treatment, 20% the first session only because no other videotape was available, and 18% some other configuration. For Phase 2 sets, 54% contained 3 consecutive sessions, 21% contained 2 consecutive sessions, 21% contained 1 session only, and 4% contained some other configuration. In MDFT, 245 sessions were selected from 74 cases. Due to dropout, 34% of cases had Phase 1 tapes only. Across the 245 sessions, 51% were Phase 1 tapes, 29% were Phase 2 tapes between sessions 6 and 12, and 20% were Phase 2 tapes between 13 and 25. For Phase 1 sets, 57% contained the first 2 sessions of treatment, 15% the first session only, and 28% some other configuration. For Phase 2 sets, 67% contained 3 consecutive sessions, 19% contained 2 consecutive sessions, 4% contained 1 session only, and 10% contained some other configuration. A total of 14% of sessions were with the adolescent alone, 12% with parent(s) alone, and 74% conjointly with the adolescent and parent(s).

VTAS–R. The current study utilized VTAS–R mean ratings from Hogue, Dauber, Faw, et al. (2006). For CBT, only therapist–adolescent alliance was coded. For MDFT, judges completed separate alliance protocols while viewing the tape, one for the adolescent if present and one for the parent (or two, then averaged) if present. Due to resource limitations, only 1 session apiece from Phase 1 (session 2 for 69% of cases) and Phase 2 (randomly selected) were coded. In CBT, a total of 71 sessions (42 Phase 1 and 29 Phase 2) across 47 cases were coded for adolescent alliance. In MDFT, 73 sessions were rated for adolescent alliance (47 Phase 1 and 26 Phase 2) and 72 sessions for parent alliance (48 Phase 1 and 24 Phase 2) across 67 cases; the total number of MDFT sessions coded was 93 (58 Phase 1 and 35 Phase 2); most MDFT tapes (n = 52, or 56%) received both adolescent and parent ratings because both members participated in the session for at least 20 minutes.

Observational Coding Procedures

The study utilized three separate groups of observational coders. Seven TBRS–C judges for CBT were recruited from a private mental health clinic specializing in cognitive–behavioral therapy; CBT judges had an average of 4.8 years (SD = 3.7) postgraduate therapy experience and 4.0 years (SD = 3.4) experience in CBT. Eight TBRS–C judges for MDFT were recruited from a community mental health clinic specializing in family therapy; MDFT judges had an average of 6.1 years (SD = 8.3) postgraduate therapy experience and 4.9 years (SD = 7.9) experience in family therapy. The five VTAS–R judges were psychology graduate students. Training and rating procedures for all three coding

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

groups were equivalent. Judges were trained in group format during weekly meetings over 4 months and demonstrated acceptable mean reliability (ICC > .65) on a preponderance of items before coding study tapes. During coding, each coding group met biweekly to enhance training and prevent rater drift. Two judges were randomly paired to each case (TBRS–C) or tape (VTAS–R).

Data Analysis

This original clinical trial used a longitudinal panel design in which participants were assessed at baseline, discharge, and 6-month follow-up. Individual client change was analyzed by using latent growth curve modeling (LGC; Duncan & Duncan, 2004). LGC produces growth curve estimates for each individual and aggregates individual trajectories to estimate mean growth parameters (intercept and slope), characterizing the sample in terms of the average baseline value of the dependent measure (i.e., intercept) and the rate and shape of change over time (i.e., slope). Missing data were addressed with full information maximum likelihood estimation, which produces unbiased parameter estimates under the assumption that data are missing at random (Schafer & Graham, 2002).

LGC proceeded in three stages. First, a series of growth curve models was tested to determine the overall shape of the individual change trajectories for the five outcome variables. This was done to determine whether symptom improvement in the study sample ($N = 136$) was comparable with that demonstrated in the full trial sample ($N = 224$; Liddle et al., in press). Based on results from the original trial for 6-month follow-up data, two forms of growth were examined: no change and linear change. Second, therapist–adolescent alliance scores were added as a covariate to control for therapeutic relationship and general process factors. Third, the fidelity variables (adherence, competence), treatment condition (CBT vs. MDFT), and their interaction terms were added to the models to examine the impact of fidelity on symptom change over time by testing the significance of the slope growth parameter. To maximize power of the statistical models for the modest sample size, fidelity–outcome relations were assessed in a single model that included participants from both conditions (excepting tests involving marijuana use frequency, for reasons described below). This across-condition analytic approach generated more model stability (i.e., fewer convergence problems) than did the within-condition approach, and it allowed for inferences about generic fidelity effects that might be shared by individual and family-based treatments for adolescent substance use. Also, post hoc within-condition analyses were conducted whenever there was an interaction involving treatment condition.

Product interaction terms, representing the interaction between fidelity variable and treatment condition, were created following procedures outlined by Aiken and West (1991) and Curran, Bauer, and Willoughby (2004), in which the two variables were grand mean centered and then multiplied. Significant two-way interactions were probed by investigating the relation between fidelity and outcome separately within each treatment condition. In these probing procedures, we examined simple slopes for all combinations of low, mean, and high values (high and low were defined as ± 1 standard deviation) for each fidelity variable to determine the overall outcome trend (Aiken & West, 1991). Three-way interactions were also created to represent the interaction among adherence, competence, and treatment condition; however, we found no significant effects for these higher order terms.

After conducting tests of linear fidelity–outcome relations, we examined curvilinear adherence–outcome relations by using the squared values of the centered adherence scores ("adherence^2"), following the work of Barber et al. (2006). As recommended by Aiken and West (1991), we included the linear adherence term in these models along with the adherence^2 variable. We also included interaction terms involving adherence^2 and treatment condition because analyses for probing this interaction were easily conducted (testing the adherence^2 effect within each condition) and interpretable. We did not test the adherence^2 by competence interaction term or any three-way interaction terms because methods for probing interactions involving curvilinear effects with two continuous variables have not been fully developed for LGC models (Curran et al., 2004).

Growth curve modeling was conducted with Mplus software Version 4.1 (Muthén & Muthén, 1998–2007). We tested fidelity–outcome effects by using pseudo-z tests (coeff./std. error > 1.96) of the slope parameter. To control for nesting effects, in all analytic models we used the sandwich variance estimator (Diggle, Heagerty, Liang, & Zeger, 2002) available in Mplus. The sandwich estimator produces corrected standard errors in the presence of non-independent data due to nested data structures, in this case, clients nested within therapists.[1]

Results

Preliminary Analyses

Therapist main effects for process and outcome variables. We tested all study variables for therapist main effects, which refer to mean-level differences among multiple therapists in a given study with respect to implementing treatments or producing outcomes (Crits-Christoph & Mintz, 1991). First, we examined therapist differences in adherence, competence, and alliance in separate analyses of variance for each condition: Therapist was entered as a fixed-factor independent variable and the process variables as the dependent variables. A significant effect was found only for adherence in CBT, $F(5, 58) = 4.58$, $p < .01$. Post hoc Scheffe tests identified a mean difference between Therapist 2 ($M = 2.51$, $SD = 0.38$) and Therapist 4 ($M = 2.03$, $SD = 0.28$). When six outlier cases (3 *SD* above or below the mean) belonging to these therapists were removed from therapist main effects analyses, results were no longer significant. To conserve sample size, these cases were retained in fidelity–outcome analyses. We then performed analyses of covariance for each outcome variable within each condition: Therapist was entered as a fixed-factor independent variable, pretreatment score on the outcome as a covariate, and posttreatment

[1] We used the sandwich estimator rather than multilevel modeling to control nesting effects due to the instability of random effects estimates generated from a small number of therapists (nine) at Level 2. To produce stable estimates in multilevel models, 20 or more Level 2 units are needed (Kreft, 1996; Snijders & Bosker, 1999).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 2
Linear Effects of Treatment Adherence and Therapist Competence on Change in Key Outcome Variables Across Treatment Conditions

Outcome variable	Growth factor mean		Adherence effects		Competence effects	
	Slope	*SE*	Slope	*SE*	Slope	*SE*
Drug use frequency[a]						
CBT	-0.64^{***}	0.08	-0.41^{*}	0.18	0.23	0.13
MDFT	-1.39^{**}	0.55	−1.03	1.23	−0.33	1.28
Parent-reported externalizing symptoms	-3.39^{***}	0.37	-2.07^{*}	0.87	0.53	0.55
Youth-reported externalizing symptoms	-1.67^{***}	0.37	−0.32	0.63	−0.09	0.54
Parent-reported internalizing symptoms	-1.72^{***}	0.33	0.27	0.84	−0.44	0.59
Drug use problems	-4.64^{***}	1.01	−0.31	2.37	−0.90	1.51

Note. The growth factor mean represents linear change in the outcome variable from baseline to 6 months posttreatment after adjusting for all covariates: adherence, competence, treatment condition, and interaction terms. Adherence effects represent the association of adherence with the outcome variable, and competence effects represent the association of competence with the outcome variable. Intercept effects were excluded from the table to simplify the presentation of study findings; these data are available from the authors on request.
[a] For this outcome only, models were fit separately for multidimensional family therapy (MDFT) and cognitive–behavioral therapy (CBT) to achieve adequate fit.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

score (discharge, follow-up) as the dependent variable. No significant effects were found in either condition.

Distributions and growth trajectories of outcome variables. We examined the distributions of all five outcome variables at each timepoint to determine whether they were approximately normal, which is a basic assumption of maximum likelihood estimation implemented in LGC modeling. The distributions for marijuana use frequency showed significant skew and kurtosis; as a result, we transformed this variable by using a natural log function, making the distributions approximately normal. We then examined the linear growth models for each outcome, as described above. As seen in Table 2, all outcome variables but marijuana use frequency demonstrated a significant linear slope estimate for the growth factor mean, indicating that the problem behavior significantly decreased over time. Because a linear growth model did not fit for marijuana use with the combined sample, we split the sample to identify what kind of growth model fit each condition separately. For CBT, a linear growth curve model was fit by using a continuous variable consisting of log-transformed values. For MDFT, a linear multinomial logistic model was fit by using a dichotomized variable in which 0 represented no use and 1 represented one or more occasions of use. We therefore split the sample when conducting all fidelity–outcome analyses involving this outcome.

Linear Fidelity–Outcome Relationships

Main effects: Adherence and competence. There were significant adherence effects for two outcomes (see Table 2). There was a main effect for adherence on marijuana use frequency in CBT (mean slope = −0.41; pseudo-z = –2.27; $p < .05$; 95% CI = 2.45, –2.09), indicating that stronger adherence predicted a greater decrease in use from baseline to 6 months posttreatment.[2] The Cohens' d effect size for this association (van Lier, Muthén, van der Sar, & Crijnen, 2004) was .44; according to Cohen (1988), d = .20 is a small effect, .50 is medium, and .80 large. No corresponding adherence effect on marijuana use frequency was found in MDFT. There was also a main effect for adherence on parent-reported externalizing symptoms (mean slope = –2.07; pseudo-z = –2.36; $p < .05$; 95% CI = –2.94, –1.18; d = .37), indicating that across CBT and MDFT, stronger adherence predicted greater decreases in externalizing. No main effects for competence were found on any outcome variable.

Interaction effects: Adherence and competence with treatment condition. There was a significant adherence by condition interaction for parent-reported internalizing symptoms (mean slope = 6.08; pseudo-z = 3.38; $p < .001$; 95% CI = 2.48, 9.68; d = 1.79). We then examined adherence effects on internalizing separately in each condition, following the post hoc probing procedures described above. Surprisingly, in CBT the relation between adherence and outcome was positive, indicating that stronger adherence was associated with increased internalizing problems (mean slope = 3.00; pseudo-z = 2.17; $p < .05$; 95% CI = 1.62, 4.38; d = .79). In contrast, in MDFT there was a marginal trend in the expected direction (mean slope = –1.88; pseudo-z = –1.67; p = .10; 95% CI = –3.01, –0.73; d = .35), with stronger adherence predicting reduced internalizing. Interactions between adherence, competence, and treatment condition were non-significant for all other outcomes.

Curvilinear Adherence–Outcome Effects

Testing curvilinear adherence effects. All fidelity–outcome analyses were reconducted with the addition of the quadratic adherence term, adherence^2, and its interaction term with treatment condition. As in the linear effects analyses, the quadratic adherence term was tested separately in CBT and MDFT for the marijuana use frequency dependent variable only; these within-condition analyses failed to converge (likely due to more complex models having additional predictors) and did not approach significance. For analyses involving the remaining outcomes, one curvilinear effect was found: There was a main effect for adherence^2

[2] To be consistent with MDFT, these analyses of marijuana use frequency in CBT were reconducted by using a dichotomized outcome variable; results were identical to those using a continuous variable with log-transformed values, that is, a significant main effect for adherence.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

on parent-reported internalizing symptoms across both conditions (mean slope = –1.50; pseudo-z = –2.46; $p < .05$; 95% CI = –2.11, –0.89; d = .40).

To interpret curvilinear effects, it is helpful to graph the function being tested. We used equations provided by Curran et al. (2004) for constructing curvilinear slopes and inserted the parameter estimates derived from the LGC test of the quadratic adherence–internalizing relation. The graph was plotted by substituting raw adherence scores for the centered scores (following Aiken & West, 1991) because centered scores yield values that are less than zero and render the graph difficult to interpret. The equations and resulting curves are presented in Figure 1. The identified relation between adherence and internalizing is a U-shaped function—similar to that reported by Barber et al. (2006)—suggesting that moderate levels of adherence predicted the best outcomes (i.e., lowest internalizing scores) over time, whereas low and high levels of adherence predicted relatively worse internalizing scores.

Reinterpreting linear adherence effects. It is necessary to reinterpret the linear effect of adherence on internalizing symptoms in light of the curvilinear effect of adherence on the same outcome. The contradictory findings reported above for the linear tests—stronger adherence predicted worse outcome in CBT but better outcome in MDFT—can now be regarded as an incomplete picture of the underlying curvilinear relation between adherence and internalizing that characterizes both treatment conditions. Statistical tests modeling a straight-line function cannot adequately capture the more complex U-shaped function modeled by curvilinear tests. In this case, the significant curvilinear test essentially resolves the apparent paradox presented by the linear tests: It is simultaneously true that both too little adherence (captured by the MDFT linear effect) and too much adherence (captured by the CBT linear effect) predicted worse outcome for internalizing problems.

Discussion

This study found that treatment adherence predicted treatment outcome in manualized behavioral interventions for substance abuse and related behavior problems in urban adolescents. Treatment adherence was linked to improvement in multiple outcomes up to 6 months after discharge. Adherence promoted therapeutic change across two different outpatient approaches: individual cognitive–behavioral therapy and multidimensional family therapy. In CBT, greater levels of adherence predicted greater declines in marijuana use. In both CBT and MDFT, stronger adherence predicted greater reductions in parent reports of externalizing behaviors. Also in both conditions, intermediate levels of adherence predicted the largest declines in parent reports of internalizing behaviors, with high and low adherence predicting smaller improvements—a curvilinear (or quadratic) effect on internalizing. Adherence–outcome effects were small-to-medium in size. Contrary to hypotheses, therapist competence was not related to any outcome in either condition, nor did it moderate the impact of adherence on outcome.

These findings support the contention that treatment adherence plays an important role in the success of empirically based behavioral interventions for adolescent mental health problems. Adherence–outcome studies have generated inconsistent results over the past 2 decades, with some studies reporting favorable adherence effects, others no effects, and still others iatrogenic effects. Several explanations for negative results have been of-

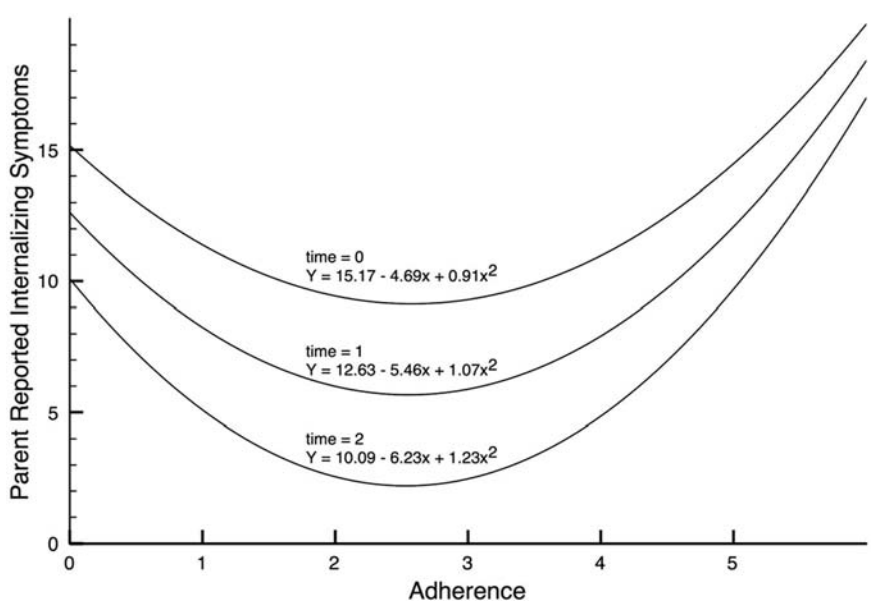

Figure 1. Curvilinear effect of treatment adherence on parent-reported internalizing symptoms across treatment conditions. Times of 0, 1, and 2 represent adherence–internalizing relations at intake, discharge, and 6-month follow-up, respectively. Equations are those reported by Curran et al. (2004, Equation 19), incorporating values derived from growth curve parameter estimates generated by the latent growth curve models that examined the curvilinear effect of adherence (i.e., adherence^2) on internalizing symptoms, after controlling for the linear effect of adherence.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

fered, including measurement insensitivity, low mean adherence levels and restricted ranges, lack of differentiation from other process variables such as therapeutic alliance, and the inability of adherence measures to account for flexible deviations from treatment protocols that prove beneficial for some sessions or clients (Miller & Binder, 2002; Perepletchikova & Kazdin, 2005). The current study was well positioned to detect adherence effects by (a) examining two treatment models with track records of adherence, differentiation, and treatment technique–outcome links (Hogue et al., 1998, 2004); (b) measuring fidelity at the level of molar treatment modules that are broadly applicable across sessions and clients; and (c) utilizing an instrument with adequate reliability and construct validity. Also, previous research with substance using and delinquent adolescents has documented similar positive adherence effects (Huey et al., 2000; Schoenwald, Sheidow, Letourneau, & Liao, 2003), raising the possibility that adherence may be particularly salient to manualized treatments for youths with problem behaviors.

This study is one of the first to replicate the innovative work of Barber et al. (2006) with regard to curvilinear adherence effects. Barber et al. (2006) suggested that inconsistent findings for linear adherence effects—does greater adherence predict better outcome?—might be due to the existence of underlying curvilinear adherence–outcome relations. Our results support this conjecture. There were contradictory findings for linear effects on internalizing symptoms: Stronger adherence predicted decreased symptoms in MDFT but increased symptoms in CBT. However, when higher order, curvilinear effects were examined, the linear relations faded to non-significance and a quadratic relation emerged across conditions: Intermediate adherence was associated with the greatest improvement, while weaker and stronger levels produced less improvement. Like Barber et al. (2006), we interpret curvilinear effects to be a caution against being too lax or too strict in adhering to treatment protocols.

Unlike Barber et al. (2006), who found linear adherence–outcome relations for a single outcome only, we found linear effects for marijuana drug use and externalizing symptoms in addition to curvilinear effects for internalizing symptoms. Based on these results, it appears that strong adherence does not unilaterally signify inflexible model implementation that sabotages treatment strength. This begs the question: Under what conditions does (over)extensive use of prescribed therapy processes diminish treatment gains, such that treatment adherence should be tempered by other considerations? For the current study, at least one mechanism seems plausible. A sizable portion of adolescent drug users has co-occurring anxiety and mood disorders (see Rowe, Liddle, Greenbaum, & Henderson, 2004). For this subgroup, the featured elements of manualized treatments designed to target drug use and externalizing behaviors specifically may need to be moderated in favor of auxiliary interventions that directly target internalizing problems: changing unrealistic negative thoughts, interpersonal problem-solving skills, relaxation training, and so forth (Compton, Burns, Egger, & Robertson, 2002; Weisz, McCarty, & Valeri, 2006). Further research is required to verify the prevalence of curvilinear adherence effects for internalizing problems in teen drug-using populations and, if confirmed, to illuminate the mechanisms of action.

It was surprising that therapist competence bore no relation to treatment outcome, nor did it influence the relation between adherence and outcome. If evidence of weak or null competence effects continues to accumulate in equal measure with evidence of small positive effects in randomized trials (Barber, Sharpless, Klostermann, & McCarthy, 2007), this will challenge clinical researchers to prove rather than presume that greater competence begets better outcome. The newly intriguing question "Does therapist competence matter?" has at least three affirmative responses: (a) Yes, if it can be correctly measured, which is to say, account for contextual variables such as intervention timing and appropriateness, responsiveness to clients, and adaptability across sessions and cases (Elkin, 1999; Miller & Binder 2002; Waltz et al., 1993). The current study took significant steps toward contextual assessment of therapist competence by utilizing expert judges, sampling multiple consecutive sessions in early and later treatment, and incorporating aspects of therapist skill and responsiveness into the competence coding system (Barber et al., 2007). The derived competence scores for CBT and MDFT had adequate distributions but only fair-to-weak interrater reliabilities. (b) Yes, but only up to a point. Therapists need to meet an acceptable standard of competent model delivery akin to a "red line" benchmark (Shaw & Dobson, 1988), but beyond that, scaling to greater heights of observed competence may not translate into greater clinical success. Such hypotheses about how competence relates to outcome—the shape of the competence–outcome curve—can now be readily examined with random regression and growth curve modeling techniques. (c) Yes, but primarily so in routine clinical settings with front-line providers exhibiting a wide range of therapy skills. To date, competence research has been conducted almost exclusively in controlled conditions with research-trained therapists, which narrows the band of fidelity scores and potentially mutes fidelity–outcome relations (Dobson & Singer, 2005).

This study has several strengths that instill confidence in the reliability and generalizability of findings. Participants included an ethnically diverse group of adolescents and families from a large urban area. Parallel fidelity measures were used for both CBT and MDFT, which permitted us to combine all participants into a single analysis to increase power and generalizability (Elkin, 1999). Interrater reliability was robust for the adherence items in both scales even though ratings covered molar-level therapy modules rather than discrete treatment techniques (e.g., Barber, Liese, & Abrams, 2003; Morgenstern et al., 2001). Fidelity was measured across multiple sessions for each case, and adherence and competence impacts were analyzed after controlling for therapeutic alliance, which reduces "third variable" confounds in the form of non-specific processes and therapeutic relationship factors (Perepletchikova & Kazdin, 2005). However, there was not enough power to examine fidelity–alliance interactions, which have been found in several other studies (e.g., Barber et al., 2006).

One significant limitation of the study was the low interrater reliability of competence scales. Reliability for individual TBRS–C items was generally weak and well below the magnitude found for competence items on most discrete techniques scales (e.g., Barber et al., 2003). Reliabilities of the averaged competence ratings (.56 for CBT, .55 for MDFT) were modest but in keeping with the magnitude of competence ratings in some studies (e.g., Barber & Crits-Christoph, 1996; James, Blackburn, Milne, & Reichfelt, 2001), though decidedly lower than in others (e.g., Carroll et al., 2000; Moyers, Martin, Manuel, Hendrickson, & Miller, 2005). Barber et al. (2007) noted that interrater reliability

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

estimates derived from competence measures used in controlled trials tend to be low and list several possible explanations, including differences in how much attention judges pay to different aspects of treatment delivery, difficulties in operationalizing competence, and the use of uniformly competent therapists in randomized studies (which dampens the total variance in competency scores). For the current study, null findings for competence effects must be considered tentative in light of the subpar ICCs obtained.

Note also that interrater reliabilities in this study were estimated by using the one-way analysis of variance model, ICC(1,2; Shrout & Fleiss, 1979). This model is a conservative approach when used in sampling designs with missing judge data, that is, when every judge does not rate every target or when a balanced incomplete block design (Fleiss, 1981) is not achieved. Although newer methods for calculating ICCs are available that account for missing data via maximum likelihood estimation (Konishi & Shimizu, 1994), they are rarely used in observational coding studies. Such methods yield more precise ICCs under conditions of missing data because they accurately estimate variance components for both target and judge, and they can accommodate nested sampling designs (e.g., sessions nested within therapist and treatment condition; see Barber, Foltz, Crits-Christoph, & Chittams, 2004).

Another study limitation is that a full examination of treatment dropout effects on fidelity–outcome relations (see Hedeker & Gibbons, 1997) was beyond the scope of this study. Also, by utilizing case-level fidelity scores that were averaged across individual sessions, we precluded the possibility of examining change in fidelity during the course of treatment. Improvement in fidelity across sessions and cases is thought to be evidence of a "learning curve" in therapist training studies (Crits-Christoph et al., 1998), and these trends may also meaningfully impact client outcomes in routine practice. Finally, the use of observational coding methods, while adding the rigor of more objective assessment, presents limitations as well. Judges who do not observe (most) every session are not able to track the clinical progress of the case across treatment, which hampers their ability to provide fully informed, case-specific assessments of competence (Barber et al., 2007; Waltz et al. 1993). An alternative strategy is to collect therapist-report and/or supervisor-report fidelity data on most or all sessions, pending further verification that fidelity instruments are reliable and valid when used as self-report tools by front-line supervisors (adherence and competence) as well as therapists and clients (adherence only; Carroll, Nich, & Rounsaville, 1998; Henggeler et al., 1997, 1999; Schoenwald et al., 2004).

As efforts to move research-based treatments into everyday clinical settings gather steam, treatment fidelity concerns will remain at the forefront of dissemination research. The multifaceted relation between treatment adherence and client outcome found in this study awaits replication and further definition in the lab and in the field with adolescent substance-using samples and with additional age groups, ethnic groups, mental health disorders, and treatment models. Based on current findings, future research on adherence–outcome links should routinely explore both linear and curvilinear effects. This will help mark the path for clinicians treading the fine line between protocol adherence and flexible deviation to serve the needs of individual clients and clinical subgroups.

References

Achenbach, T. M. (1991a). *Manual for the Child Behavior Checklist/4–18 and 1991 Profile.* Burlington: University of Vermont.

Achenbach, T. M. (1991b). *Manual for the Youth Self-Report and 1991 profile.* Burlington: University of Vermont Department of Psychiatry.

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Newbury Park, CA: Sage.

Barber, J. P., & Crits-Christoph, P. (1996). Development of a therapist adherence/competence rating scale for supportive–expressive dynamic psychotherapy: A preliminary report. *Psychotherapy Research, 6,* 81–94.

Barber, J. P., Crits-Christoph, P., & Luborsky, L. (1996). Effects of therapist adherence and competence on patient outcome in brief dynamic therapy. *Journal of Consulting and Clinical Psychology, 64,* 619–622.

Barber, J. P., Foltz, C., Crits-Christoph, P., & Chittams, J. (2004). Therapists' adherence and competence and treatment discrimination in the NIDA Cocaine Collaborative Treatment Study. *Journal of Clinical Psychology, 60,* 29–41.

Barber, J. P., Gallop, R., Crits-Christoph, P., Frank, A., Thase, M., Weiss, R. D., & Gibbons, M. (2006). The role of therapist adherence, therapist competence, and alliance in predicting outcome of individual drug counseling: Results from the National Institute on Drug Abuse Collaborative Cocaine Treatment Study. *Psychotherapy Research, 16,* 229–240.

Barber, J. P., Liese, B. S., & Abrams, M. J. (2003). Development of the Cognitive Therapy Adherence and Competence Scale. *Psychotherapy Research, 13,* 205–221.

Barber, J. P., Sharpless, B., Klostermann, S., & McCarthy, K. S. (2007). Assessing intervention competence and its relation to therapy outcome: A selected review derived from the outcome literature. *Professional Psychology: Research and Practice, 38,* 493–500.

Breslin, C., Sobell, L. C., & Sobell, M. B. (1996). Aftercare telephone contacts with problem drinkers can serve a clinical and research function. *Addiction, 91,* 1359–1364.

Carroll, K. M., Connors, G. J., Cooney, N. L., DiClemente, C. C., Donovan, D. M., Kadden, R. R., et al. (1998). Internal validity of Project MATCH treatments: Discriminability and integrity. *Journal of Consulting and Clinical Psychology, 66,* 290–303.

Carroll, K. M., Kadden, R. M., Donovan, D. M., Zweben, A., & Rounsaville, B. J. (1994). Implementing treatment and protecting the validity of the independent variable in treatment matching studies. *Journal of Studies in Alcohol, 12,* 149–155.

Carroll, K. M., Nich, C., & Rounsaville, B. J. (1998). Utility therapist session checklists to monitor delivery of coping skills treatment for cocaine abusers. *Psychotherapy Research, 8,* 307–320.

Castonguay, L. G., Goldfried, M. R., Wiser, S., Raue, P. J., & Hayes, A. M. (1996). Predicting the effect of cognitive therapy for depression: A study of unique and common factors. *Journal of Consulting and Clinical Psychology, 64,* 497–504.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6,* 284–290.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Compton, S. N., Burns, B. J., Egger, H. E., & Robertson, E. (2002). Review of the evidence base for treatment of childhood psychopathology: Internalizing disorders. *Journal of Consulting and Clinical Psychology, 70,* 1240–1266.

Crits-Christoph, P., & Mintz, J. (1991). Implications of therapist effects for the design and analysis of comparative studies of psychotherapies. *Journal of Consulting and Clinical Psychology, 59*(1), 20–26.

Crits-Christoph, P., Siqueland, L., Blaine, J., Frank, A., Luborsky, L., Onken, L. S., et al. (1999). Psychosocial treatments for cocaine depen-

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

dence: National Institute on Drug Abuse Collaborative Cocaine Treatment Study. *Archives of General Psychiatry, 56,* 493–502.

Crits-Christoph, P., Siqueland, L., Chittams, J., Barber, J. P., Beck, A. T., Frank, A., et al. (1998). Training in cognitive, supportive–expressive, and drug counseling therapies for cocaine dependence. *Journal of Consulting and Clinical Psychology, 66,* 484–492.

Curran, P. J., Bauer, D. J., & Willoughby, M. T. (2004). Testing main effects and interactions in latent curve analysis. *Psychological Methods, 9,* 220–237.

Dennis, M. L., Godley, S. H., Diamond, G., Tims, F. M., Babor, T., Donaldson, J., et al. (2004). The Cannabis Youth Treatment (CYT) study: Main findings from two randomized trials. *Journal of Substance Abuse Treatment, 27,* 197–213.

Diamond, G. S., & Diamond, G. M. (2002). Studying a matrix of change mechanisms: An agenda for family-based process research. In H. A. Liddle, D. A. Santisteban, R. F. Levant, & J. H. Bray (Eds.), *Family psychology: Science-based interventions.* Washington, DC: American Psychological Association.

Diggle, P. J., Heagerty, P., Liang, K. Y., & Zeger, S. L. (2002). *Analysis of longitudinal data.* New York: Oxford University Press.

Dobson, K. S., & Singer, A. R. (2005). Definitional and practical issues in the assessment of treatment integrity. *Clinical Psychology: Science and Practice, 12,* 384–387.

Duncan, T. E., & Duncan, S. C. (2004). An introduction to latent growth curve modeling. *Behavior Therapy, 35,* 333–363.

Elkin, I. (1999). A major dilemma in psychotherapy outcome research: Disentangling therapists from therapies. *Clinical Psychology: Science and Practice, 6,* 10–32.

Fals-Stewart, W., O'Farrell, T. J., Freitas, T. T., McFarlin, S. K., & Rutigliano, P. (2000). The timeline follow-back reports of psychoactive substance use by drug-abusing patients: Psychometric properties. *Journal of Consulting and Clinical Psychology, 68,* 134–144.

Feeley, M., DeRubeis, R. J., & Gelfand, L. A. (1999). The temporal relation of adherence and alliance to symptom change in cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 67,* 578–582.

Fleiss, J. L. (1981). Balanced incomplete block designs for inter-rater reliability studies. *Applied Psychological Measurement, 5,* 105–112.

Frank, E., Kupfer, D. J., Wagner, E. F., McEachran, A. B., & Cornes, C. (1991). Efficacy of interpersonal psychotherapy as a maintenance treatment of recurrent depression: Contributing factors. *Archives of General Psychiatry, 48,* 1053–1059.

Hartley, D. E., & Strupp, H. H. (1983). The therapeutic alliance: Its relationship to outcome in brief psychotherapy. In J. Masling (Ed.), *Empirical studies of psychoanalytical theories* (Vol. 1, pp. 1,038). Hillsdale, NJ: Analytical Press.

Hedeker, D., & Gibbons, R. D. (1997). Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychological Methods, 2,* 64–78.

Henggeler, S. W., Melton, G. B., Brondino, M. J., Scherer, D. B., & Hanley, J. H. (1997). Multisystemic therapy with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination. *Journal of Consulting and Clinical Psychology, 65,* 821–833.

Henggeler, S. W., Pickrel, S. G., & Brondino, M. J. (1999). Multisystemic treatment of substance abusing and dependent delinquents: Outcomes, treatment fidelity, and transportability. *Mental Health Services Research, 1,* 171–184.

Henry, W. P., Strupp, H. H., Butler, S. F., Schacht, T. E., & Binder, J. L. (1993). Effects of training in time-limited dynamic psychotherapy: Changes in therapist behavior. *Journal of Consulting and Clinical Psychology, 61,* 434–440.

Hogue, A., Dauber, S., Chinchilla, P., Fried, A., Henderson, C. E., Inclan, J., et al. (in press). Assessing fidelity in individual and family therapy for adolescent substance abuse. *Journal of Substance Abuse Treatment.*

Hogue, A., Dauber, S., Faw, L., Cecero, J. J., & Liddle, H. A. (2006). Early therapeutic alliance and treatment outcome in individual and family therapy for adolescent behavior problems. *Journal of Consulting and Clinical Psychology, 74,* 121–129.

Hogue, A., Dauber, S., Samuolis, J., & Liddle, H. A. (2006). Treatment techniques and outcomes in multidimensional family therapy for adolescent behavior problems. *Journal of Family Psychology, 20,* 535–543.

Hogue, A., Liddle, H. A., Dauber, S., & Samuolis, J. (2004). Linking session focus to treatment outcome in evidence-based treatments for adolescent substance abuse. *Psychotherapy: Theory, Research, Practice, & Training, 41,* 83–96.

Hogue, A., Liddle, H. A., Rowe, C., Turner, R. M., Dakof, G. A., & LaPann, K. (1998). Treatment adherence and differentiation in individual versus family therapy for adolescent substance abuse. *Journal of Counseling Psychology, 45,* 104–114.

Huey, S. J., Henggeler, S. W., Brondino, M. J., & Pickrel, S. G. (2000). Mechanisms of change in multisystemic therapy: Reducing delinquent behavior through therapist adherence and improved family and peer functioning. *Journal of Consulting and Clinical Psychology, 68,* 451–467.

James, I. A., Blackburn, I. M., Milne, D. L., & Reichfelt, F. K. (2001). Moderators of trainee therapists' competence in cognitive therapy. *British Journal of Clinical Psychology, 40,* 131–141.

Konishi, S., & Shimizu, K. (1994). Maximum likelihood estimation of an intraclass correlation in a bivariate normal distribution with missing observations. *Communication in Statistics—Theory and Methods, 23,* 1593–1604.

Kreft, I. G. G. (1996). *Are multilevel techniques necessary? An overview, including simulation studies.* Los Angeles, CA: California State University.

Liddle, H. A. (2002). *Multidimensional family therapy for adolescent cannabis users: Cannabis Youth Treatment Series, Volume 5* (DHHS Pub. No. 02–3660). Rockville, MD: Center for Substance Abuse Treatment, Substance Abuse and Mental Health Services Administration.

Liddle, H. A., Dakof, G. A., Parker, K., Diamond, G., Barrett, K., & Tejeda, M. (2001). Multidimensional family therapy for adolescent drug abuse: Results of a randomized clinical trial. *American Journal of Drug and Alcohol Abuse, 27,* 651–688.

Liddle, H. A., Dakof, G. A., Turner, R. M., Henderson, C. E., & Greenbaum, P. E. (in press). Treating adolescent drug abuse: A randomized trial comparing Multidimensional Family Therapy and Cognitive Behavior Therapy. *Addiction.*

Liddle, H. A., Rowe, C. L., Dakof, G. A., Ungaro, R. A., & Henderson, C. E. (2004). Early intervention for adolescent substance abuse: Pretreatment to posttreatment outcomes of a randomized controlled trial comparing multidimensional family therapy and peer group treatment. *Journal of Psychoactive Drugs, 36,* 49–63.

Miller, S. J., & Binder, J. L. (2002). The effects of manual-based training on treatment fidelity and outcome: A review of the literature on adult individual psychotherapy. *Psychotherapy: Theory, Research, Practice, and Training, 39,* 184–198.

Morgenstern, J., Morgan, T. J., McCrady, B. S., Keller, D. S., & Carroll, K. M. (2001). Manual guided cognitive-behavioral therapy training: A promising method for disseminating empirically supported substance abuse treatments to the practice community. *Psychology of Addictive Behaviors, 15,* 83–88.

Moyers, T. B., Martin, T., Manuel, J. K., Hendrickson, S. M. L., & Miller, W. R. (2005). Assessing competence in the use of motivational interviewing. *Journal of Substance Abuse Treatment, 28,* 19–26.

Muthén, L., & Muthén, B. (1998–2007). *Mplus user's guide.* Los Angeles, CA: Author.

Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice, 12,* 365–383.

Rowe, C. L., Liddle, H. A., Greenbaum, P. E., & Henderson, C. E. (2004). Impact of psychiatric comorbidity on treatment of adolescent drug abusers. *Journal of Substance Abuse Treatment, 26,* 129–140.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7,* 147–177.

Schoenwald, S. K., Sheidow, A. J., Letourneau, E. J., & Liao, J. G. (2003). Transportability of multisystemic therapy: Evidence for multilevel influences. *Mental Health Services Research, 5,* 223–239.

Shaw, B. F., & Dobson, K. S. (1988). Competency judgments in the training and evaluation of psychotherapists. *Journal of Consulting and Clinical Psychology, 56,* 666–672.

Shaw, B. F., Elkin, I., Yamaguchi, J., Olmstead, M., Vallis, T. M., Dobson, K. S., et al. (1999). Therapist competence ratings in relation to clinical outcome in cognitive therapy of depression. *Journal of Consulting and Clinical Psychology, 67,* 837–846.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428.

Snijders, T., & Bosker, R. (1999). *Multilevel analysis.* London: Sage.

Sobell, L. C., & Sobell, M. B. (1996). *Timeline followback user's guide: A calendar method for assessing alcohol & drug use.* Toronto, Ontario, Canada: Addiction Research Foundation.

Startup, M., & Shapiro, D. A. (1993). Therapist treatment fidelity in prescriptive vs. exploratory psychotherapy. *British Journal of Clinical Psychology, 32,* 443–456.

Stiles, W. B., Honos-Webb, L., & Surko, M. (1998). Responsiveness in psychotherapy. *Clinical Psychology: Science and Practice, 5,* 439–458.

Turner, R. M. (1992). Launching cognitive-behavioral therapy for adolescent depression & drug abuse. In S. Budman, M. Hoyt, & S. Friedman (Eds.), *Casebook of brief therapy* (pp. 135–156). New York: Guilford.

van Lier, P. A. C., Muthén, B., van der Sar, R. M., & Crijnen, A. A. M. (2004). Preventing disruptive behavior in elementary schoolchildren: Impact of a universal classroom-based intervention. *Journal of Consulting and Clinical Psychology, 72,* 467–478.

Waldron, H. B., & Kaminer, Y. (2004). On the learning curve: The emerging evidence supporting cognitive-behavioral therapies for adolescent substance abuse. *Addiction, 99*(Suppl. 2), 93–105.

Waldron, H. B., Slesnick, N., Brody, J. L., Turner, C. W., & Peterson, T. R. (2001). Treatment outcomes for adolescent substance abuse at 4- and 7-month assessments. *Journal of Consulting and Clinical Psychology, 69,* 802–813.

Waltz, J., Addis, M. E., Koerner, K., & Jacobson, N. S. (1993). Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology, 61,* 620–630.

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.

Winters, K. C., Latimer, W. W., & Stinchfield, R. D. (2002). Clinical issues in the assessment of adolescent alcohol and other drug use. *Behavior Research & Therapy, 40,* 1443–1456.

Received June 19, 2007
Revision received March 17, 2008
Accepted March 19, 2008 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.