2016

Evaluating Measures of Fidelity for Substance Abuse Group Treatment With Incarcerated Adolescents

Shayna S. Bassett
University of Rhode Island, sbassett@uri.edu

L.A.R. Stein
University of Rhode Island, larstein@uri.edu

See next page for additional authors

Citation/Publisher Attribution
Bassett, S. S., Stein, L. A. R., Rossi, J. S., & Martin, R. A. (2016). Evaluating Measures of Fidelity for Substance Abuse Group Treatment With Incarcerated Adolescents. *Journal of Substance Abuse Treatment, 66*, 9-15. doi: 10.1016/j.jsat.2016.02.011
Available at: https://doi.org/10.1016/j.jsat.2016.02.011

HHS Public Access

Author manuscript

J Subst Abuse Treat. Author manuscript; available in PMC 2017 July 01.

Published in final edited form as:

J Subst Abuse Treat. 2016 July ; 66: 9–15. doi:10.1016/j.jsat.2016.02.011.

Evaluating Measures of Fidelity for Substance Abuse Group Treatment with Incarcerated Adolescents

Shayna S. Bassett, M.S.[a], **L. A. R. Stein, Ph.D.**[a,b,c], **Joseph S. Rossi, Ph.D.**[a], and **Rosemarie A. Martin, Ph.D.**[c]

[a]Department of Psychology, University of Rhode Island, Kingston, RI 02881

[b]The Rhode Island Training School, Cranston, RI 02920, USA

[c]Center for Alcohol and Addiction Studies, Brown University, Providence, RI 02906, USA

Abstract

The evaluation of treatment fidelity has become increasingly important as the demand for evidence-based practice grows. The purpose of the present study is to describe the psychometric properties of two measures of treatment fidelity that can be used by therapists and supervisors -one for group-based Cognitive-Behavioral Therapy (CBT) and one for combined Substance Education and Twelve-Step Introduction (SET) for adolescent substance use. At the end of group sessions (CBT $n = 307$; SET $n = 279$), therapists and supervisors completed an evaluation measure assessing adherence to certain core components of the intervention. The supervisor version of the fidelity measure also included items for rating the level of competency the therapist demonstrated when providing each component of the intervention. Results from split-half cross-validation analyses provide strong support for an 11-item, three-factor CBT fidelity measure. Somewhat less consistent but adequate support for a nine-item, two-factor SET fidelity measure was found. Internal consistencies ranged from acceptable to good for both the CBT and SET adherence scales and from acceptable to good for the CBT and SET competency scales, with the exception of the CBT Practices competency scale. Preliminary validation of the measures suggests that both measures have adequate to strong factor structure, reliability, and concurrent and discriminant validity. The results of this study have implications for research and clinical settings, including the supervision process.

Keywords

Treatment fidelity; substance use treatment; group counseling; adolescents

Correspondence should be sent to: Shayna S. Bassett, M.S., Department of Psychology, Social Sciences Research Center, University of Rhode Island, 130 Flagg Rd., Kingston, RI 02881, Phone: 401-874-4261; Fax: 401-874-5562; ; Email: SSoenksen@my.uri.edu

1. Introduction

The evaluation of treatment fidelity has become increasingly important as the demand for evidence-based practice grows. Fidelity monitoring has become a requirement within efficacy research on treatment interventions and has quickly become a principle focus of treatment dissemination. One of the best ways to replicate the success of interventions obtained in research settings is to maintain high implementation fidelity when transporting to community settings (Carroll et al., 2007). Fidelity to empirically based treatments is positively linked to more optimal outcomes for clients in community-based settings (Barber et al., 2006; Houge, Henderson et al, 2008).

Treatment fidelity, also referred to as treatment integrity, is defined as the extent to which an intervention is delivered as intended by the protocol. Three components of treatment delivery that should be monitored to assess treatment fidelity include adherence, competence, and treatment differentiation (Waltz, Addis, Koerner, & Jacobsen, 1993). Treatment adherence refers to the degree to which the therapist employs procedures prescribed by the intervention and avoids proscribed procedures (Breitenstein et al., 2010; Houge, Henderson et al, 2008; Perepletchikova, Treat, & Kazdin, 2007). Treatment competence encompasses the level of skill the therapist demonstrates in implementing the prescribed procedures. Competence is concerned with *how well* the protocol is implemented (Breitenstein et al., 2010). Lastly, treatment differentiation determines whether treatments differ from each other along critical dimensions (Perepletchikova & Kazdin, 2005; Perepletchikova, Treat, & Kazdin, 2007).

In spite of recommendations to develop rigorous tools that allow for reliable and valid data on treatment fidelity to be collected (Carroll et al., 2000; Perepletchikova & Kazdin, 2005), surprisingly few studies sufficiently assess for it. According to results from a review conducted by Perepletchikova, Treat, & Kazdin (2007), only 3.5% of studies investigating psychosocial interventions adequately assessed treatment fidelity. Furthermore, most fidelity rating instruments have yet to establish psychometric soundness (Baer et al., 2007).

Relatively few measures have been developed to assess the treatment fidelity of evidence-based interventions for adolescent substance use (e.g., Hogue, Dauber et al., 2008; Resko, Walton, Chermack, Blow, & Cunningham, 2012). Therefore, additional evaluation tools for assessing fidelity to treatments that are both feasible and cost-effective are needed (Carroll et al., 2000; Perepletchikova, Treat, & Kazdin, 2007).

Which method should be used to evaluate therapist adherence and competence in clinical practice remains in question. Reasonable agreement has been found between therapist self-ratings and independent ratings of skill acquisition following training in Motivational Interviewing (MI; Hartzler, Baer, Dunn, Rosengren, & Wells, 2007). Martino, Ball, Nich, Frankforter, & Carroll (2009) also found that therapists and supervisors generally agreed on whether strategies fundamental to Motivational Enhancement Therapy occurred during sessions. This led the authors to suggest that the use of therapists' self-report may be a reasonable and cost-effective way to determine therapists' use of basic counseling techniques. However, poor correspondence was found between supervisor and therapist

ratings for more complex counseling techniques. Moreover, therapists have been found to overestimate their level of adherence to evidence-based interventions compared to supervisor ratings (Carroll, Martino, & Rounsaville, 2010; Carroll et al., 2000; Martino et al., 2009). As a result, therapist self-report should not be relied on alone to monitor treatment fidelity to empirically based treatments (Carroll et al., 2010); however, it may act as a cue to therapists regarding important aspects of intervention. For example, Hartzler et al. (2007) found therapist self-ratings were associated with more effective use of MI and increased therapist self-awareness.

Treatment fidelity measures also play an important role in clinical supervision. Using a session rating form, supervisors are able to review sessions and provide therapists with valuable feedback regarding their relative strengths and weakness (Sampl & Kadden, 2001). Furthermore, if both the therapist and the supervisor complete a fidelity measure at the end of a treatment session, the two measures can be compared, allowing for any notable differences in the interpretation of the session to be discussed.

Substance use treatments in correctional facilities are frequently provided in group format and often include components of Cognitive-Behavioral Therapy (CBT) and Twelve-Step approaches (The Correctional Association of New York, 2011). Twelve-step facilitation was found to be used at least sometimes by 74% of facilities (Substance Abuse and Mental Health Services Administration [SAMHSA], 2013). A national survey of substance use treatment for juvenile offenders found that 51% of juvenile correctional facilities offered cognitive and behavioral interventions, 89% provided substance education, and 93% provided treatment in group format (Young, Dembo, & Henderson, 2007). Group-based CBT for adolescent substance use, based on the Cannabis Youth Treatment (CYT) Study (Dennis et al., 2004), has been widely disseminated with the treatment manuals freely available from SAMHSA (e.g., http://store.samhsa.gov/shin/content//SMA08-3954/SMA08-3954.pdf). Although quality control procedures were provided for CBT in the CYT Study, the fidelity rating form used in the project and included in the treatment manual was never formally validated.

As noted above, group-based substance education and Twelve-Step approaches are also widely utilized for adolescents and at relatively low cost (Center for Substance Abuse Treatment, 1999). A treatment manual for group-based Substance Education and Twelve-Step Introduction (SET) for adolescent substance use was created for a randomized clinical trial described in Section 2.1; however, the fidelity rating form included in this treatment manual has yet to be formally validated. Having validated fidelity measures for these treatment approaches (i.e., CBT and SET) would be of great assistance to professionals working with substance-involved youth, particularly given how frequently these interventions are available, and is consistent with the movement toward disseminating evidence-based interventions. Importantly, such measures would also facilitate the supervision process.

The purpose of the present study is to describe the psychometric properties of two measures of treatment fidelity for adolescent substance use that can be used by therapists and supervisors - one for group-based Cognitive-Behavioral Therapy (CBT) and one for group-

based Substance Education and Twelve-Step Introduction (SET). This description focuses on measure composition, reliability, and concurrent and discriminant validity.

2. Material and methods

2.1 Overview of the Clinical Trial

These data were collected as part of a randomized clinical trial comparing two group treatments for substance abusing incarcerated adolescents (R01 DA-13375; PI-Stein). Participants were recruited at a state juvenile correctional facility in the Northeast. Immediately after adjudication, adolescents were identified as potential candidates for the study if they were between the ages of 14 and 19 years, inclusive, and sentenced to the facility for between 4 and 12 months, inclusive. Consent was obtained from legal guardians, and assent was obtained from adolescents (adolescents 18 years or older provided consent). Adolescents were included in the study if they met any of the following substance use criteria: (a) they used marijuana or drank at least monthly, or binge drank (>5 standard drinks for boys, >4 for girls) at least once in the year before incarceration; (b) they used marijuana or drank in the 4 weeks before the offense for which they were incarcerated; or (c) they used marijuana or drank in the 4 weeks before they were incarcerated. Enrollment in substance use programming did not require for the study, nor for the juvenile correctional facility, that youth have a substance use disorder.

Of the 1,280 adolescents screened for the study, 205 met screening criteria and completed the consent procedure. Of the baseline sample, 38.8% were of Hispanic ethnicity and racial groups were as follows: 36.8% African American, 30.9% White, 7.9% Native American, 5.3% Pacific Islander, 4.6% Asian American, and 14.5% self-identified as other. Most were boys (88.2%), average age was 16.9 years (SD= 1.09), and the average number of previous times detained or incarcerated was 2.54 (SD= 2.41). Percent using any alcohol or marijuana in the three months before incarceration was 81.0% and 90.7%, respectively. Percent binge-drinking in the three months before incarceration was 67.8%. In the year prior to incarceration, 27.7% and 59.9% qualified for alcohol and marijuana dependence, respectively. All procedures that were utilized received Institutional Review Board approval.

Following baseline assessment, adolescents were randomized to two sessions of individually delivered Motivational Interviewing (MI; Stein & Clair, 2010a) or two sessions of combined Meditation-Relaxation Training (RT; Stein & Clair, 2010b). Following MI, adolescents received 10 group-based sessions of CBT (see Stein, 2005) modeled after the CYT manuals (Sampl & Kadden, 2001; Webb, Scudder, Kaminer, & Kadden, 2002). Similarly, following RT, adolescents received 10 group-based sessions of SET (see Rose et al., 2005) which were created based on the standard content used by the juvenile correctional facility. As individual treatments are not relevant to the current study, they will not be further discussed. Both group interventions were manualized, including fidelity procedures. MI and RT strategies could be utilized during group CBT and SET sessions, respectively, although they were not a focus of the group interventions. Group sessions lasted approximately 75 minutes each and occurred 1–3 times per week with about 3 participants per group session. On average, adolescents received 8 group sessions over 6.5 weeks. Of the 586 total groups conducted

over the course of the study, 307 were CBT and 279 were SET. Groups were gender-segregated and rolling admissions was used.

2.2 Development of Fidelity Measures

2.21 CBT Measure - Therapist/Supervisor—The 31-item CBT fidelity measure was adapted from the Supervisor Group Session Rating form used in the CYT Study (Sampl & Kadden, 2001; Webb et al., 2002). Items from the original measure were retained. A few items were modified to be more applicable to incarcerated versus community youth (e.g., "To what extent did the therapist assess clients' *desire to use* marijuana, alcohol, or other substances since the last session," vs. "To what extent did the therapist assess clients' *use* of marijuana or other substances since the last session"). Eight additional items were created and added to the measure to assess proscribed interventions and treatment differentiation. For the CBT measure, proscribed interventions or behaviors refer to those that should be unique to SET (e.g., "To what extent did the therapist emphasize an abstinence-only orientation?"). Proscribed interventions were rated according to adherence only (rating procedures are described below in Section 2.23).

2.22 SET Measure - Therapist/Supervisor—The 26-item SET fidelity measure was modeled after the CBT fidelity measure and was included in the original SET manual. Fidelity items thought to capture the interventions and behaviors specific to SET were generated based on content in the treatment manual for this intervention condition (e.g., “To what extent did you encourage participants to maintain abstinence after release?"). In addition, some items were common across both intervention conditions since some procedures should occur in both treatments (e.g., "To what extent did you manage behaviors that were disruptive to the group process [aggression, war stories, excessive profanity]?"; “To what extent was it difficult to engage the group?”). The study principle investigator generated about 30 items which were then reviewed by the investigative team for appropriateness and clarity. The 26 items analyzed in this study were those remaining. The final item on the measure assesses the extent to which 10 proscribed CBT approaches were used (e.g., practice of skills, assertiveness, identification of triggers, etc.) with ratings ranging from 1 (not at all; 0–1 of 10) to 5 (extensively; 8–10 of 10) and provides an estimate of treatment differentiation. This item assessed adherence only.

2.23 Format of the ratings—Therapists completed either the CBT fidelity measure or the SET fidelity measure at the end of a group session, depending upon which type of therapy was provided, to assess adherence to core components of the intervention in response to the stem “to what extent did you…” rated on a Likert scale ranging from 1 (not at all) to 5 (extensively). Supervisors assessed the same CBT and SET adherence items in addition to rating the competency, or level of skill, the therapist demonstrated when providing each component of the intervention. A Likert scale ranging from 1 (very poor) to 7 (excellent) was used to assess competency with an additional 0 (not done) response option.

2.3 Group Report Grid

Therapists completed the Group Report Grid, which was obtained and adapted from the original CYT manual (Sampl & Kadden, 2001), at the end of every group session to monitor

elements of the group as a whole. The elements monitored with this measure include quantity of participation (0 = none to 3 = high), quality of participation (0 = none to 3 = high), interest in substance use (1 = none to 3 = interest in both alcohol and other drugs), clinical status throughout group (1 = poor to 5 = excellent), and disruptive behavior (aggressive, interrupts, profanity, sexually inappropriate, glorifying drug use, etc.).

2.4 Rater Training

Therapists had approximately 160 hours of manualized training (Rose, Klein, Stein, Lebeau-Craven, & Justus, 2005; Stein, 2005) which included reading manuals, question/answer, seeing demonstrations of sessions, and role-plays with feedback. During training sessions, therapists completed fidelity measures which were compared with supervisor fidelity ratings for agreement and discussion if ratings did not agree within +/−1 point on the Likert scale. Based on fidelity procedures in the CYT Study, feedback was provided on a particular session until fidelity criteria were reached. In brief, meeting fidelity criteria involved reaching a skill rating of at least adequate on items that were to occur in the session. Prior to working with youth, therapists were required to pass a dress-rehearsal with confederates. Therapists completed the fidelity measure directly following the group session, endorsing one appropriate response per item. Therapists received one hour of group supervision per week and one hour of individual supervision per month. During supervision, cases and/or sessions were reviewed, questions answered, and role-plays to enhance practice were utilized as needed. Videos were also reviewed as desired to enhance practice. Therapists were one man and five women; all six were Caucasian; four had an MA, one had a BS and one had a PhD. Each therapist conducted both intervention types.

All group sessions were video-recorded for fidelity coding by supervisors. Supervisors provided ratings during or following the viewing of a video-recorded session. Supervisors could stop or back up the video recording while completing the ratings. Two PhD-level clinical investigators with expertise in behavioral intervention for substance abusing adolescents made supervision and fidelity determinations for both intervention types (see recommendations in manuals).

As specified in the CYT Study, once therapists reached fidelity criteria (see definition above), supervisors occasionally reviewed to maintain performance standards. Any therapist falling below fidelity criteria received added individual supervision and monitoring until fidelity criteria were again reached. In addition, for the parent study, we added that the proscribed treatment was to occur at a minimum or not at all (ratings of 1 or 2 on the adherence scale) as rated by the supervisor. In the event that ratings of proscribed activities did not meet this standard, again, therapists received added individual supervision and monitoring until proscribed activities occurred at a minimum or not at all. Sessions were randomly chosen for supervisor ratings, and feedback was provided immediately during the next supervision session. During the course of the study, supervisors rated 30% of group sessions to monitor fidelity, and of those, 41% were randomly selected and double-coded for inter-rater reliability. Experts trained in the interventions provided double ratings for supervisor forms. This was performed throughout the study (approximately every 4 months)

to prevent drift. Disagreements were discussed among raters to resolve discrepancies and maintain practice and quality procedures.

2.5 Analyses

For the 586 groups that were conducted, the study produced 307 and 85 completed CBT fidelity measures for therapists and supervisors, respectively; and 279 and 88 SET fidelity measures for therapists and supervisors, respectively. Therapist and supervisor ratings were combined for analyses of adherence items. When both therapist and supervisor ratings were available for a single group session, only the supervisor rating was included in analyses to maintain independence of observations. The split-half cross-validation technique was used to allow for stronger conclusions to be made regarding the reliability and validity of the measures (Redding, Maddock, & Rossi, 2006). First, half of the measures of each version (CBT and SET) were randomly selected for use in exploratory analyses, which began with item analyses for each version. Two items were immediately removed from the CBT measure and one item from the SET measure due to low variance. The adherence ratings for remaining items on each version were entered into principal components analyses (PCA) using parallel analysis (Horn, 1965) to determine the number of components to extract and varimax rotation to determine item loadings (Redding et al., 2006; Velicer, Eaton, & Fava, 2000). After the dimensional structure was identified for each measure, items that failed to load ≥ .40 on any factor or that loaded ≥ .40 on two or more factors were removed. PCAs were again conducted until all remaining items loaded uniquely at .40 or more on their respective target factors. Item-scale correlations and item contributions to scale internal consistency, using coefficient alpha, were also examined to identify items for deletion. Next, confirmatory factor analyses (CFA) with correlated factors were conducted using the second (independent) half of each sample to test the factor structure identified in the exploratory analyses for each measure. Goodness of fit was evaluated using the following criteria: Root mean square error of approximation (RMSEA); standardized root mean square residual (SRMR); and comparative fit index (CFI). Values closer to zero indicate better fit for RMSEA and SRMR and closer to 1.0 for CFI (Kline, 2011).

Using the complete sample for each measure, correlations were obtained between the identified adherence and competence scales to provide a measure of concurrent validity. To provide additional support for concurrent validity, correlations were obtained between the CBT and SET adherence and competence scales and other measures of group behavior as assessed using the Group Report Grid. Because only supervisors rated competency, correlations examining the competency scales were computed using only the supervisor ratings for each measure.

Intra-class correlation coefficients (ICCs) were calculated for double-coded sessions to provide an estimate of rater reliability. All ICCs were calculated with α set at .05 using a 2-way random model so that results could be generalized to other samples and raters. Consistency between raters was believed to be an appropriate bar to set over absolute agreement; as a result Type = Consistency was chosen. Lastly, Average Measures was chosen to capture the reliability of the mean of raters.

Items assessing proscribed interventions on both measures were remarkably skewed and could not be corrected by transformation. As a result, formal analyses could not be conducted to determine how well identified scales discriminated between interventions. However, mean ratings on CBT-related and SET-related scales during CBT intervention, collected using the CBT fidelity measure, were compared to assess discriminant validity with the expectation that CBT-related scales would be rated higher. Similarly, mean ratings on SET-related and CBT-related indices during SET intervention, collected using the SET fidelity measure, were compared with the expectation that SET-related indices would be rated higher. Additional mean comparisons were made to compare CBT- and SET-related scales across intervention types. Finally, CBT adherence and competency scales were compared to SET adherence and competency scales using independent *t*-tests to determine whether the two treatments were conducted equally well in the parent study.

3. Results

3.1 Factor Structure

A series of PCAs produced a final exploratory factor structure of three factors for the CBT fidelity measure (CBT Practices, MI Practices, and Group Engagement) consisting of 11 items accounting for 67.0% of the variance. Similarly, two components were identified for the SET fidelity measure (SET Practices and Group Engagement). Nine items composed the scales accounting for 67.3% of the variance. The resulting factor loadings for each item are presented in Table 1 for the CBT Fidelity Measure and Table 2 for the SET Fidelity Measure.

Results of the CFAs on the factor structures identified in the exploratory PCAs demonstrate the three-factor structure of the CBT measure was largely confirmed (^{2} [41, N = 155] = 100.38, p < .01; CFI = .92; RMSEA = .10, 90% CI = [.07, .12]; SRMR = .08). All three factors were positively and significantly correlated (CBT Practices and MI Practices: r = .29; CBT Practices and Group Engagement: r = 27; MI Practices and Group Engagement: r = . 26). While some indices indicated adequate goodness of fit, less consistent support was found for the two-factor structure of the SET measure (χ^2 [26, N = 140] = 95.55, p < .001; CFI = .88; RMSEA = .14, 90% CI = [.11, .17]; SRMR = .09). Both factors were positively and significantly correlated (SET Practices and Group Engagement: r = .35). The resulting factor loadings for each item are presented in Table 1 for the CBT Fidelity Measure and Table 2 for the SET Fidelity Measure.

3.2 Evaluation of Reliability

Descriptive statistics for the full sample for all scales are presented in Table 3. Internal consistencies were obtained for each scale of each measure. The internal consistencies of the CBT adherence scales were excellent (α = .91) for CBT Practices, acceptable for MI Practices (α = .63), and good for Group Engagement (α= .77). The internal consistency of the CBT Practices competency scale was low (α = .40) but was good for the MI Practices competency scale (α = .74). Items comprising the Group Engagement scale did not include competency ratings. For the SET fidelity measure, internal consistencies were good for both

adherence scales (SET Practices scale: α = .87; Group Engagement scale: α = .72) and acceptable for the SET Practices competency scale (α = .65).

The following criteria were used to classify the magnitude of the ICCs for the double-coded group sessions: < .40 = *poor*; .40 – .59 = *fair*; .60 – .74 = *good*; .75 – 1.00 = *excellent* (Cicchetti, 1994). ICCs for the CBT adherence scales were as follows: CBT Practices = .73, $p < .001$ (*good*); MI Practices = .58, $p < .001$ (*fair*); Group Engagement = .64, $p < .001$ (*good*). The ICCs for the SET adherence scales were stronger: SET Practices = .64, $p < .001$ (*good*); Group Engagement = .87, $p < .001$ (*excellent*). ICCs for the competency scales were also calculated: CBT Practices = .69, $p < .001$ (*good*); MI Practices = .76, $p < .001$ (*excellent*); SET Practices = .73, $p < .001$ (*good*).

3.3 Concurrent Validity

Pearson product-moment correlations were calculated for the adherence and competence scales of the CBT and SET measures (see Table 4). The magnitude of the correlations for the CBT and SET adherence scales were low, suggesting the scales on each measure are independent of each other. Generally, the competence scales for both the CBT and SET measures were significantly and positively related with the adherence scales, suggesting that therapists in the parent study demonstrated skill in adhering to the treatment protocols. Table 5 contains correlations among the CBT and SET adherence and competency scales with other variables used to measure group behavior from the Group Report Grid. These variables include the average: 1) quantity of participation; 2) quality of participation; 3) interest in substance use; 4) clinical status; and 5) disruptive behaviors. In general, greater adherence to CBT Practices and Group Engagement as assessed by the CBT fidelity measure was associated with more optimal group behavior (i.e., higher quality and quantity of group participation, better clinical status, fewer misbehaviors); these findings were less consistent with the MI Practices adherence scale, although MI seemed to be used more when clinical status was poor. Similarly, greater adherence to SET Practices and Group Engagement as assessed by the SET fidelity measure was associated with more optimal group behavior. The competence scales from both measures were generally not found to be significantly related to group behavior as measured by the Group Report Grid, with the exception of SET Practices where more competence was associated with less interest in substance use.

3.4 Discriminant Validity

Mean comparisons were made within group treatments to illustrate that interventions consistent with CBT did not occur in SET and that interventions consistent with SET did not occur within CBT. Therapists and supervisors rating CBT group sessions rated CBT-specific elements of the protocol, as measured by the CBT Practices scale ($M = 2.56$; $SD = .45$) and the MI Practices scale ($M = 2.51$; $SD = .47$), higher than SET-specific elements ($M = 1.01$; $SD = .05$). Similarly, those rating SET group sessions rated SET interventions, as measured by the SET Practices scale ($M = 2.32$; $SD = .40$), higher than the CBT-specific index ($M = 1.04$; $SD = .21$). Furthermore, CBT-specific elements were rated higher in CBT groups ($M = 2.56$; $SD = .45$) than in SET groups ($M = 1.04$; $SD = .21$), and SET-specific interventions were rated higher in SET groups ($M = 2.32$; $SD = .40$) compared to CBT groups ($M = 1.01$; $SD = .05$).

Additional comparisons were formally made to determine whether level of fidelity differed between treatment groups. Therapists were found to adhere to SET Practices significantly less than CBT Practices ($t[583] = 6.79$, $p < .001$, Cohen's d [d] = .56) and MI Practices ($t[584] = 5.24$, $p < .001$, $d = .43$). No difference was found between the treatment groups on the Group Engagement scale ($t[583] = .42$, $p = .68$, $d = .03$). SET Practices were conducted as competently as CBT Practices ($t[171] = .19$, $p = .85$, $d = .03$) but more competently than MI Practices ($t[171] = 2.45$, $p = .02$, $d = .37$).

4. Discussion

This study examined the psychometric properties of treatment fidelity measures for CBT and SET group interventions as assessed by therapists and supervisors. Results from the split-half cross-validation analyses provide strong support for the 11-item, three-factor CBT fidelity measure. Somewhat less consistent but adequate support for the nine-item, two-factor SET fidelity measure was found. The factor analytic procedure used allows us to draw firmer conclusions regarding the reliability and validity of the measures (Redding et al., 2006).

Internal consistencies ranged from acceptable to good for both the CBT (.63 – .91) and SET adherence scales (.72 – .87) and from acceptable to good for the CBT and SET competency scales (.65 – .74), with the exception of the CBT Practices competency scale which was quite low (.40). Therefore, findings from this scale should be interpreted with caution. In general, these results provide strong support regarding the reliability of the measures.

Correlations between adherence and competency scales were in the expected direction (that is, overall, higher adherence ratings correlated with higher levels of competence, and lower adherence ratings correlated with lower levels of competence) and provide support for good convergent validity. Interestingly, more competence in SET Practices related to less adherence to Group Engagement strategies, perhaps because SET was designed to be more didactic. The scales on both measures performed well when correlated with other measures of group behavior assessed at the same time. In general, greater fidelity to treatment was associated with better quantity and quality of participation and less interest in substance use among group members. The Group Engagement scale on each measure was negatively correlated with misbehaviors during group sessions. This suggests that adolescents who are engaged in treatment are less likely to act out during group sessions. Higher MI adherence ratings correlated with poorer clinical status, suggesting that therapists utilized MI skills in particular when youth were in distress. Only adherence and competence on SET Practices was related to reduced interest in substance use. The emphasis in SET was on knowing the negative effects of drugs and being abstinent. In contrast, the emphasis in CBT was on learning skills in order to reduce harm, which could include abstinence.

Both measures were found to demonstrate good discriminant validity. This was evidenced during CBT group sessions by therapists and supervisors rating CBT-specific elements more highly than SET elements, and SET-specific elements more highly than CBT elements during SET group sessions. Furthermore, SET interventions in SET were rated more highly

than SET interventions in CBT, and CBT interventions in CBT were rated more highly than CBT interventions in SET.

Comparisons of the two treatment groups indicated that therapists adhered to the CBT protocol significantly more than the SET protocol. No differences between the two treatment groups were found in regards to Group Engagement or the CBT and SET Practices competency scales. Therapists were found to perform SET Practices more competently than MI Practices. Although these statistical differences were found, the differences appear less clinically meaningful. For example, CBT Practices, SET Practices, and Group Engagement adherence scale means fell between 2 (*a little*) and 3 (*somewhat*) on the Likert scale used. Similarly, SET Practices, MI Practices, and CBT Practices competency scale means fell between 4 (*adequate*) and 5 (*good*). These ratings are consistent with the expectations outlined by the group treatment manuals used in this study (Rose et al., 2005; Stein, 2005) and the CYT Study (Sample & Kadden, 2001; Webb et al, 2001), and suggest that both interventions were delivered with adequate or better fidelity in the parent study.

The current study is not without limitations. Because a gold standard does not yet exist for assessing treatment fidelity for CBT or SET group treatment, convergent validity could not be determined. The SET measure did not include items assessing for the individual treatment (i.e., RT) adolescents received but the CBT measure did (i.e., MI). This is a limitation of the parent study but may suggest these fidelity measures are more generalizable to other treatment settings given the frequency with which MI/CBT is implemented while RT/SET is more unique to this study. Similarly, the inclusion of more than one item to assess proscribed behaviors in SET may have been more advantageous. However, the item that was added assessed the presence of CBT components comprehensively and was designed to detect variation in the use of CBT. While it is not possible to track which CBT-specific behaviors might have occurred using this measure, it is possible to track how much CBT-specific behaviors occurred. Most fidelity measures of which we are aware do not even make an effort to distinguish between two intervention types while implementing a specific intervention. That we added at least one comprehensive CBT item to the SET fidelity measure was beyond what is usually done on fidelity tools. PCAs could not be conducted on the competency scales for either measure due to limited sample size and were assumed to be similar in structure as the adherence scales. Additional research is needed to formally establish the structure of these scales. The psychometric properties resulting from the current study are based on a relatively small sample and in one clinical context (i.e., juvenile correctional facility), possibly limiting the generalizability of the findings. However, given the need for treatment and volume of youth in correctional settings, perhaps it is important that we conducted the study in this setting. Lastly, Group Grid ratings were based on therapist reports, as compared to more objective ratings of group behavior, such as third party blind and independent observers.

4.1 Conclusions and Implications

The results of this study have implications for research and clinical settings. First, given how few studies have examined psychometric properties of fidelity rating tools (Baer et al., 2007), this study begins to address this gap in the literature by providing psychometric

analyses of two fidelity measures that can be used to assess competence and adherence to manualized group treatment interventions. Second, the use of these treatment fidelity measures in research protocols may allow investigators to reduce random error and have more confidence in the results of their studies (Bellg et al., 2004). Third, these measures can be used by therapists and supervisors in order to tap various vantage points. This may be clinically useful for the supervision process, during which therapists and supervisors can compare session rating measures to determine points of agreement and differences. Furthermore, these measures contain scales that have been created to target constructs relevant to each treatment as compared to therapists and supervisors discussing item-level ratings. In addition, the scales are in common between therapist and supervisor versions so that both can compare their adherence assessment of the session, which can facilitate discussion and learning. Implementing this practice is likely to increase the therapist's awareness of CBT and SET skills and improve his/her ability to self-monitor (Sampl & Kadden, 2001). Further validation of these measures in other studies may suggest they are appropriate for use in a variety of research studies and clinical settings with adolescent substance users. Future research may determine if adherence and/or competency measures are related to outcomes or mediate outcomes on substance use.

Acknowledgments

This research was supported in part by NIH Grants R01 DA-018851 (National Institute on Drug Abuse, PI-Stein) and G20RR030883 (National Center for Research Resources). NIH had no role in the study design, data collection, analysis, interpretation, manuscript writing, or the decision to submit the paper for publication.

References

Baer JS, Ball SA, Campbell BK, Miele GM, Schoener EP, Tracy K. Training and fidelity monitoring of behavioral interventions in multi-site addictions research: A review. Drug and Alcohol Dependence. 2007; 16(87):107–118. [PubMed: 17023123]

Barber JP, Gallop R, Crits-Christoph P, Frank A, Thase ME, Weiss RD, Gibbons MBC. The role of therapist adherence, therapist competence, and alliance in predicting outcome of individual drug counseling: Results from the National Institute Drug Abuse Collaborative Cocaine Treatment Study. Psychotherapy Research. 2006; 16(4):229–240.

Bellg AJ, Borrelli B, Resnick B, Hecht J, Minicucci DS, Ory M, Czajkowski S. Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the NIH behavior change consortium. Health Psychology. 2004; 23(5):443–451. [PubMed: 15367063]

Breitenstein SM, Gross D, Garvey C, Hill C, Fogg L, Resnick B. Implementation fidelity in community-based interventions. Research in Nursing and Health. 2010; 33(2):164–173. [PubMed: 20198637]

Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. Implementation Science. 2007; 2(40)

Carroll KM, Martino S, Rounsaville BJ. No train, no gain? Clinical Psychology: Science and Practice. 2010; 17(1):36–40.

Carroll KM, Nich C, Sifry RL, 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]

Center for Substance Abuse Treatment. Treatment of Adolescents with Substance Use Disorders; Treatment Improvement Protocol (TIP) Series, No. 32. Rockville, MD: Substance Abuse and Mental Health Services Administration; 1999.

Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment. 1994; 6:284–290.

Dennis M, Godley SH, Diamond G, Tims FM, Babor T, Donaldson J, Funk R. The Cannabis Youth Treatment (CYT) Study: Main findings from two randomized trials. Journal of Substance Abuse Treatment. 2004; 27:197–213. [PubMed: 15501373]

Hartzler B, Baer JS, Dunn C, Rosengren DB, Wells E. What is seen through the looking glass: The impact of training on practitioner self-rating of Motivational Interviewing skills. Behavioural and Cognitive Psychotherapy. 2007; 35(4):431–445.

Hogue A, Dauber S, Chinchilla P, Friend A, Henderson C, Inclan J, Reiner RH, Liddled HA. Assessing fidelity in individual and family therapy for adolescent substance abuse. Journal of Substance Abuse Treatment. 2008; 35:137–147. [PubMed: 17997268]

Hogue A, Henderson CE, Dauber S, Barajas PC, Fried A, Liddle HA. Treatment adherence, competence, and outcome in individual and family therapy for adolescent behavior problems. Journal of Consulting and Clinical Psychology. 2008; 76(4):544–555. [PubMed: 18665684]

Horn JI. A rationale and test for the number of factors in factor analysis. Psychometrika. 1965; 30:179–185. [PubMed: 14306381]

Kline, RB. Principles and practice of structural equation modeling. 3rd. New York: The Guilford Press, NY; 2011.

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(2):181–193. [PubMed: 19396649]

Perepletchikova F, Kazdin AE. Treatment integrity and therapeutic change: Issues and research recommendations. Clinical Psychology: Science and Practice. 2005; 12(4):365–383.

Perepletchikova F, Treat TA, Kazdin AE. Treatment integrity in psychotherapy research: Analysis of the studies and examination of associated factors. Journal of Consulting and Clinical Psychology. 2007; 75(6):829–841. [PubMed: 18085901]

Redding CA, Maddock JE, Rossi JS. The sequential approach to measurement of health behavior constructs: Issues in selecting and developing measures. California Journal of Health Promotion. 2006; 4(1):83–101.

Resko SM, Walton MA, Chermack ST, Blow FC, Cunningham RM. Therapist competence and treatment adherence for a brief intervention addressing alcohol and violence among adolescents. Journal of Substance Abuse Treatment. 2012; 42:429–437. [PubMed: 22119182]

Rose, M.; Klein, J.; Stein, LAR.; Lebeau-Craven, R.; Justus, A. Substance education treatment: A manual for treating incarcerated adolescents. Rockville, MD: Brown University, Providence, RI; NIDA/NIAAA; 2005. Unpublished manual

Sampl, S.; Kadden, R. Motivational enhancement therapy and cognitive behavioral therapy for adolescent cannabis users: 5 sessions (Cannabis Youth Treatment [CYT] Series. Vol. 1. Rockville, MD: Center for Substance Abuse Treatment, Substance Abuse and Mental Health Services Administration; 2001.

Stein, LAR. Cognitive behavior therapy for alcohol and marijuana use: A manual for treating incarcerated teens. Rockville, MD: Brown University, Providence, RI; NIDA/NIAAA; 2005. Unpublished manual

Stein, LAR.; Clair, M. Motivational Enhancement therapy for alcohol and marijuana use: A manual for treating incarcerated teens. Kingston, RI: University of Rhode Island; 2010a. Unpublished manual

Stein, LAR.; Clair, M. Relaxation therapy for alcohol and marijuana use: A manual for treating incarcerated teens. Kingston, RI: University of Rhode Island; 2010b. Unpublished manual

Substance Abuse Mental Health Services Administration. National Survey of Substance Abuse Treatment Services (N-SSATS): 2012. Data on Substance Abuse Treatment Facilities. BHSIS Series S-66, HHS Publication No. (SMA) 14–4809. 2013

Rockville, MD. Treatment Behind Bars: Substance Abuse Treatment in New York Prison. New York: The Correctional Association of New York; 2011. Substance Abuse and Mental Health Services Administration. The Correctional Association of New York.

Velicer, WF.; Eaton, CA.; Fava, JL. Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In: Goffin, RD.; Helmes, E., editors. Problems and solutions in human assessment: Honoring Douglas Jackson at seventy. Boston: Kluwer; 2000. p. 41-71.

Waltz J, Addis ME, Koerner K, Jacobsen NS. Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. Journal of Consulting and Clinical Psychology. 1993; 61:620–630. [PubMed: 8370857]

Webb, C.; Scudder, M.; Kaminer, Y.; Kadden, R. The motivational enhancement therapy and cognitive behavioral therapy supplement: 7 sessions of cognitive behavioral therapy for adolescent cannabis users (Cannabis Youth Treatment [CYT] Series, Vol. 2; DHHS Pub. No. SMA 02-3659). Rockville, MD: Center for Substance Abuse Treatment, Substance Abuse and Mental Health Services Administration; 2002.

Young DW, Dembo R, Henderson CE. A national survey of substance abuse treatment for juvenile offenders. Journal of Substance Abuse Treatment. 2007; 32(3):255–266. [PubMed: 17383550]

Highlights

- Strong support was found for an 11-item, three-factor CBT fidelity measure.
- Adequate support for a nine-item, two-factor SET fidelity measure was found.
- Validation suggests both measures have adequate to strong psychometric properties.
- These can be used to assess competence and adherence to group treatment protocols.

Table 1

CBT fidelity form items and factor loadings

CBT Fidelity Form Factors	PCA Factor Loading	CFA Factor Loading
CBT Practices		
To what extent did you discuss or address the participants' current commitment to abstinence or reduced use?	.87	.89
To what extent did you assess the participants' desire to use alcohol, marijuana, or other substances since the last session?	.83	.92
To what extent did you provide one or more specific assignments for the participants to engage in between sessions?	.81	.62
To what extent did you emphasize the importance of real life practice of skills between sessions?	.80	.66
To what extent did you respond to the participants with empathy, warmth, and acceptance?	.73	.72
MI Practices		
To what extent did you attempt to elicit self-motivational statements from the participants?	.86	.72
To what extent did you attempt to focus on the participants' ambivalence about changing their level of marijuana/alcohol use?	.80	.75
To what extent did you discuss any high-risk situations the participants have encountered and explore any coping skills used?	.67	.39
Group Engagement		
Rate the quantity of participation of this group.	.87	.89
Rate the quality of participation of this group in terms of: a) thoughtfulness; b) appropriateness; and c) effort of group members.	.87	.83
To what extent was it difficult to engage the group?	.82	.65

Table 2

SET fidelity form items and factor loadings

SET Fidelity Form Factors	PCA Factor Loading	CFA Factor Loading
SET Practices		
To what extent did you discuss or address the participants' attitudes towards abstinence?	.88	.76
To what extent did you assess the participants' interest in using alcohol, marijuana, or other substances since the last session?	.86	.73
To what extent did you encourage participants to maintain abstinence after release?	.84	.76
To what extent did you discourage continued use of alcohol, marijuana, and other drugs after release?	.81	.75
To what extent did you discuss any desire of participants to maintain abstinence after release?	.80	.82
To what extent did you discuss or address the participants' attitudes towards any of the following: a) how alcohol and marijuana work; b) negative physical, psychological, and social effects of alcohol and marijuana; c) inter- and intrapersonal context of marijuana and alcohol use (communication, self-esteem, anger); d) risky situations; e) resources after release (for example, AA/NA).	.55	.51
Group Engagement		
Rate the quantity of participation of this group.	.91	.80
Rate the quality of participation of this group in terms of: a) thoughtfulness; b) appropriateness; and c) effort of group members.	.88	.98
To what extent was it difficult to engage the group?	.73	.45

Table 3

Descriptive statistics for adherence and competency scales

	n	Mean	SD
CBT Form			
Adherence			
CBT Practices	306	2.56	.45
MI Practices	307	2.51	.47
Group Engagement	307	2.88	.59
Competency			
CBT Practices	85	4.39	.75
MI Practices	85	4.09	.85
SET Scale (proscribed)	306	1.00	.04
SET Form			
Adherence			
SET Practices	279	2.32	.40
Group Engagement	278	2.90	.57
Competency			
SET Practices	88	4.37	.64
CBT index (proscribed)	279	1.04	.21

Table 4

Correlations among CBT and SET adherence and competency scales

CBT Form				
	Adherence (n = 307)		***Competency (n = 85)***	
	MI Practices	**Group Engagement**	**CBT Practices**	**MI Practices**
Adherence Scales				
CBT Practices	.10	.09	.58***	.33**
MI Practices		.08	.34**	.51***
Group Engagement			.22*	.13
Competency Scales				
CBT Practices				.52***
MI Practices				

SET Form			
	Adherence (n = 279)		*Competency (n = 88)*
	SET Practices	Group Engagement	SET Practices
Adherence Scales			
SET Practices		.09	.38***
Group Engagement			−.22*

* p<.05,

** p<.01,

*** p<.001

Table 5

Correlations among scales and Group Grid Variables

	Average Quantity of Participation	Average Quality of Participation	Average Interest in Substance Use	Average Clinical Status	Average Misbehaviors
CBT Form					
Adherence (n = 307)					
CBT Practices	.13*	.12*	−.11	.28**	−.07
MI Practices	.03	.03	.05	−.19**	.09
Group Engagement	.72**	.73**	.00	.46**	−.42**
Competency (n = 85)					
CBT Practices	−.05	−.08	.11	−.04	−.10
MI Practices	.20	.02	.09	−.05	−.11
SET Form					
Adherence (n = 279)					
SET Practices	.21**	.18**	−.22**	.25**	.14*
Group Engagement	.60**	.48**	−.09	.34**	−.26**
Competency (n = 88)					
SET Practices	.11	.09	−.36**	.08	−.08

* = $p < .05$;

** = $p < .01$.