



Published in final edited form as:

J Subst Abuse Treat. 2016 August ; 67: 15–21. doi:10.1016/j.jsat.2016.04.006.

Toward Evidence-Based Measures of Implementation: Examining the Relationship between Implementation Outcomes and Client Outcomes

Bryan R. Garner¹, Sarah B. Hunter², Rodney R. Funk³, Beth Ann Griffin², and Susan H. Godley³

¹RTI International, P. O. Box 12194, Research Triangle Park, NC 27709-2194, USA

²RAND, P.O. Box 2138, Santa Monica, CA 90407-2138, USA

³Chestnut Health Systems, 448 Wylie Dr., Normal, IL 61761, USA

Abstract

Background—Developing consistent, valid, and efficient implementation outcome measures is necessary to advance implementation science. However, development of such measures has been limited to date, especially for validating the extent to which such measures are associated with important improvements in client outcomes. This study seeks to address this gap by developing one or more evidence-based measures of implementation (EBMIs; i.e., implementation outcome measure that is predictive of improvements in key client outcomes) for the Adolescent Community Reinforcement Approach (A-CRA), an evidence-based practice (EBP) for adolescent substance use.

Methods—Data for the current study were collected as part of a large-scale federally funded EBP dissemination and implementation initiative. The multilevel dataset included 65 substance use treatment organizations, 308 clinicians, and 5,873 adolescent clients. Adjusted multilevel regression analyses were used to examine the extent to which client-level outcome measures assessed at 6-month follow-up (i.e., substance use, emotional problems) could be predicted by four implementation outcomes: two measures of fidelity (i.e., session exposure, procedure exposure) and two measures of penetration (i.e., absolute client penetration, absolute staff penetration).

Results—Adjusting for client substance use at intake, as well as several client characteristics (e.g., age, race, criminal justice involvement), client substance use at follow-up was significantly

*Bryan R. Garner: bgarner@rti.org.
Sarah B. Hunter: shunter@rand.org
Rodney R. Funk: rfunk@chestnut.org
Beth Ann Griffin: bethg@rand.org
Susan H. Godley: sgodley@chestnut.org

Competing Interests

Authors Bryan R. Garner, Sarah B. Hunter, Rodney R. Funk, and Beth Ann Griffin declare no competing interests. Susan H. Godley was the director of the EBT Center at Chestnut Health Systems until August 2013; the EBT Center derives some revenue from training treatment provider sites in the Adolescent Community Reinforcement Approach treatment.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

lower for treatment organizations that had higher procedure exposure ($B=-1.227$, standard error [SE]=0.583, 95% confidence interval=-2.370, 0.252; $p<.05$). None of the other three implementation outcome measures were found to predict improvements in client outcomes.

Conclusions—The current study provides support for procedure exposure as an organizational-level EBMI for A-CRA. Thus, future efforts focused on implementing A-CRA could be improved by measuring and monitoring the extent to which A-CRA procedures are being delivered to clients. Additionally, given the dearth of studies that have examined the relationship between organizational-level measures of implementation and client outcomes, this article provides a prototype for future research to identify EBMI for other behavioral treatments.

Keywords

substance use; implementation outcomes; adolescent

1 Background

A plethora of evidence-based practices (EBPs)¹—those practices that have been empirically shown to be efficacious and/or effective—are available for a wide range of health conditions. Unfortunately, the difficulty of implementing EBPs in routine service settings has been documented across numerous areas of health (Institute of Medicine, 1998, 2001). The limited implementation of EBPs in routine service settings is a major issue of concern, given that hundreds of billions of dollars are spent annually to provide services that may have little (if any) evidence to support their effectiveness and given that the return-on-investment of the several hundred billions of dollars that have been spent to date developing EBPs is far from being maximized (Kerner, 2006). Implementation research (i.e., the scientific study of methods to promote the systematic uptake of research findings and other evidence-based practices into routine practice, and hence to improve the quality and effectiveness of health services; see Eccles & Mittman, 2006) has developed numerous guiding conceptual models (Klein & Sorra, 1996; Mitchell, Fisher, Hastings, Silverman, & Wallen, 2010; Simpson, 2002; Tabak, Khoong, Chambers, & Brownson, 2012; Wilson, Petticrew, Calnan, & Nazareth, 2010), but there remains an important need to develop evidence-based measures of implementation (EBMIs; i.e., implementation outcome measure with predictive validity to a distinct construct of interest measured [e.g., key client outcome] at some point in the future; see Lewis, Weiner, Stanick, & Fisher, 2015).

The concept of EBMI is relatively new, but recognition of the importance of developing implementation measures is not. For example, nearly 20 years ago, Klein and Sorra (1996) conceptualized *implementation effectiveness* (i.e., the consistency and quality of targeted organizational members' use of an innovation) as one of the earliest implementation measures. More recently, Proctor and colleagues (2011; 2009) helped advance a number of different implementation outcome measures, including (a) *acceptability*, (b) *adoption*, (c) *appropriateness*, (d) *feasibility*, (e) *fidelity*, (f) *implementation cost*, (g) *penetration*, and (h)

¹Abbreviations: A-CRA = Adolescent Community Reinforcement Approach; CSAT = Center for Substance Abuse Treatment; EBMI = evidence-based measure of implementation; EBP = evidence-based practice; FOI = Fidelity of Implementation; SAMHSA = Substance Abuse and Mental Health Services Administration; SD = standard deviation; SE = standard error.

sustainability. An even more comprehensive list of implementation measures has been described by Damschroder and colleagues (2009) as part of their Consolidated Framework for Implementation Research and by Proctor, Powell, and Feely (2014) as part of their overview of measurement in dissemination and implementation science, which together suggest the need for multiple types of implementation measures.

Implementation measures are important because they may be able to serve as key intermediate outcomes in relation to service system or clinical outcomes, which are costly and not always practical to collect. Indeed, Proctor and colleagues (2011) noted that “Once researchers have advanced consistent, valid, and efficient measures for implementation outcomes, the field will be equipped to conduct important research treating these constructs as dependent variables, in order to identify correlates or predictors of their attainment.” Implementation measures also are important because they may be able to help better understand why clinical interventions are effective (or not effective). For example, in the absence of implementation measures, if a clinical intervention is not found to be effective, it will be difficult (if not impossible) to know if this was due to shortcomings of the intervention or if the intervention was simply not implemented well. At the present time, however, EBIMs are not readily available. For example, Martinez, Lewis, and Weiner (2014) recently noted that “a paradox has emerged whereby researchers appear to be investigating implementation initiatives with instruments that may not be psychometrically sound.” These authors did not discourage the use of implementation measures without robust psychometrics because this is a necessary step toward establishing a measures psychometric quality for a given use. Nonetheless, the authors concluded that “The fact remains that without psychometrically validated instruments, investigators cannot be confident that instruments measure the purported constructs consistently.” Among several recommendations, these authors noted the need to establish instrument psychometric properties in terms of reliability and validity.

In response to the need for more psychometrically validated implementation measures, we sought to develop one or more EBIMs for the Adolescent Community Reinforcement Approach (A-CRA) (M. Dennis et al., 2004; Garner, Godley, et al., 2009; Godley et al., 2001), which is one of the most widely disseminated and implemented EBPs for adolescent substance use (Godley, Garner, Smith, Meyers, & Godley, 2011; Hunter, Ayer, Han, Garner, & Godley, 2014). A-CRA is a behavioral treatment based on a menu of 19 procedures (e.g., Happiness Scale, Goals, Communication, Problem-Solving, Caregiver Involvement), which therapists are trained to deliver during treatment sessions (Godley et al., 2001). As noted previously, research to develop EBIMs is quite limited, but the current study complements prior research by Keith and colleagues (2010), which developed and tested an organizational-level Fidelity of Implementation (FOI) measure for a nurse practitioner case management intervention. This FOI measure, which was developed using qualitative data collected from 18 staff across four medical centers, was found to be predictive of better patient outcomes regarding both patient resource utilization and patient mortality.

Implementation measures can be conceptualized, measured, and analyzed at a number of levels (e.g., organization, staff, patient/client), and the specific level or levels that are most appropriate can be debatable. That said, we elected to focus on the development of EBIMs at

the organizational-level in the current project, which we believe is justified given implementation is often a collective effort (Klein & Sorra, 1996). Additionally, while there are several potential implementation outcomes to examine, the current project focused on validating ones that were available and that included (1) fidelity and (2) penetration. In terms of fidelity, which is frequently conceptualized along one or more different dimensions (e.g., exposure, adherence, competence, participant responsiveness, program differentiation) (Dane & Schneider, 1998), we focused on exposure. More specifically, we examined *session exposure* (i.e., the number of A-CRA sessions implemented) and *procedure exposure* (i.e., the number of discrete A-CRA procedures implemented). In terms of penetration (i.e., integration of a practice within a service setting and its subsystems; Proctor et al., 2011), we examine absolute measures of *client penetration* (i.e., the number of clients receiving A-CRA) and *staff penetration* (i.e., the number of staff trained in A-CRA), which is related, yet distinct from proportional measures of penetration (proportional measure of penetration not able to be calculated as part of the current project). Given that A-CRA is an EBP used for addressing adolescent substance use, the primary client outcome of interest was improvements in substance use. We also examined, however, improvements in emotional problems because A-CRA has also been shown to help with adolescents' co-occurring emotional problems (Godley et al., 2014).

In sum, the primary goal of the current research was to develop one or more EBMI for a widely disseminated and implemented EBP for adolescent substance use (i.e., A-CRA). In general, we hypothesized greater improvements in client-level outcomes (i.e., reductions in substance use, reductions in emotional problems) among organizations with higher implementation outcome measures (i.e., session exposure, procedure exposure, client penetration, staff penetration). More specifically, because the number of A-CRA treatment sessions delivered has been shown to be an important predictor of outcome (Garner et al., 2009), we hypothesized organizations providing a greater number of A-CRA sessions on average (i.e., session exposure) would also have greater improvements in their A-CRA client's outcomes. Similarly, because the number of A-CRA treatment procedures delivered has been shown to be an important predictor of client outcomes (Garner et al., 2009), we hypothesized organizations providing a greater number of A-CRA procedures on average (i.e., procedure exposure) would also have greater improvements in their A-CRA client's outcomes. Because the absolute volume of patients has been shown to be associated with better organizational outcomes (Mesman, Westert, Berden, & Faber, 2015), we hypothesized organizations that provided A-CRA to more clients (i.e., client penetration) would also have greater improvements in their A-CRA client's outcomes. Finally, because absolute cumulative team experience has been shown to be important (Elbardissi, Duclos, Rawn, Orgill, & Carty, 2013), we hypothesized organizations with a greater cumulative A-CRA experience (i.e., staff penetration) would also have greater improvements in their A-CRA client's outcomes. In addition to providing evidence of the validity of available EBMI for A-CRA, the current research helps provide a prototype of developing EBMI for implementation research, which is limited within existing implementation research literature.

2 Methods

2.1 Data Source

Implementation data (e.g., fidelity, penetration) and client data (e.g., intake assessment, follow-up assessment) used as part of this study were collected as part of a large-scale EBP dissemination and implementation initiative funded by the Substance Abuse and Mental Health Services Administration's Center for Substance Abuse Treatment (SAMHSA/CSAT). The general goal of this initiative was to improve adolescent substance use treatment by providing multiple community-based treatment organizations with funding so that their clinical staff could learn and implement A-CRA (Godley et al., 2001), which has been shown to be effective in reducing adolescent substance use and substance-related problems (M. Dennis et al., 2004; Garner, Godley, Funk, Dennis, & Godley, 2007; Garner, Godley, et al., 2009; Godley, Godley, Dennis, Funk, & Passetti, 2007). Each treatment organization received approximately \$900,000 over a 3-year period and was able to have up to five staff participate in extensive training, feedback, and supervision in the A-CRA model at no additional cost. The training included components that have been found effective for training clinicians in EBPs, including a treatment manual, 3.5-day initial workshop, coaching/supervision sessions, and feedback on recorded sessions (Miller, Yahne, Moyers, Martinez, & Pirritano, 2004; Sholomskas et al., 2005). Godley and colleagues (2011) have described the SAMHSA/CSAT project and training model in more detail. The conduct of this project was conducted under the auspices of Institutional Review Boards of Chestnut Health Systems and the RAND Corporation.

2.2 Sample

Organizations—The dataset for the current study included 65 substance use disorder treatment organizations, which all received funding from SAMHSA/CSAT. These treatment organizations were located across 14 different states in the United States. Across the 65 organizations, the number of staff trained to deliver A-CRA ranged between 1 and 11 (mean = 4.9 [SD = 2.7]). Additionally, across the 65 organizations, the number of adolescents receiving A-CRA ranged between 39 and 565, with a median of 97 adolescents.

Staff—The dataset for the current study included 308 clinicians. Clinicians were mostly female (73%) and Caucasian (55%), with 17% Hispanic, 12% African American, and 16% mixed or other races. Clinicians ranged in age from 21 to 69, with an average age of 36.5 years (standard deviation [SD]=10.5). With regard to education, 61% had a master's degree or higher. Finally, clinicians reported an average of 4.5 years of substance use counseling experience.

Clients—The dataset for the current study included 5,873 adolescents. At the time of the intake assessment, adolescents were mostly male (75%) and had an average age in years of 16 ($SD=1.93$). In terms of race, the sample represented African American (13%), Caucasian (32%), Hispanic (36%), and mixed/other (19%). The majority of adolescents (69%) reported current criminal justice involvement. Finally, in terms of baseline use of any alcohol or other drugs, adolescents reported an average of 29 days ($SD=30.99$) of use during the past 90-day period. More specifically, adolescents reported an average of 5 days ($SD = 10.99$) of alcohol

use and an average of 26 days ($SD = 31.47$) of other substance use. The other substance use primarily reported was marijuana, with a reported average of 23 days of use ($SD = 29.45$). At 6 months, the average reported days of alcohol use was 3 days ($SD = 9.56$), and the average reported days of other substance use was 13 days ($SD = 25.38$). Again, marijuana use was the most prevalent at 6 months with an average of 12 days ($SD = 23.29$).

2.3 Measures

As noted previously, the data used were collected as part of a large SAMHSA/CSAT dissemination and implementation initiative (Godley et al., 2011). As part of this initiative, each participating treatment organization was required to use a web-based tool to document both the number of A-CRA sessions and the number of A-CRA procedures delivered to each adolescent client. More specifically, clinicians would enter each client's treatment admission and discharge dates; and after each session, they would enter the dates of the session, the A-CRA procedures delivered during the session, and the approximate number of minutes spent on each A-CRA procedure. Additionally, each organization was required to use the Global Appraisal of Individual Needs, which is a well-validated biopsychosocial assessment (Dennis, Funk, Godley, Godley, & Waldron, 2004; Dennis, Titus, White, Unsicker, & Hodgkins, 2003), for assessments of adolescents at intake and 6-month post-intake follow-up. Descriptions of the implementation measures and client outcome measures examined as part of the current study follow.

Implementation Measures—Fidelity includes several different dimensions, one of which is *exposure* and that may be represented in several ways (e.g., number of sessions, number of techniques/procedures) (Dane & Schneider, 1998). As part of the current study, we assessed fidelity via the following two measures of exposure.

Session exposure: Session exposure is an organization-level measure created for each participating organization and represents the average number of A-CRA sessions that the organization delivered to its respective adolescent clients. A typical A-CRA session is approximately an hour in length and includes at least two A-CRA procedures. The average session exposure across the 65 participating organizations was 13 ($SD=3.48$). Prior work has indicated that clinical outcomes are maximized when a client receives 7 or more sessions in combination with ten or more unique A-CRA procedures (Garner, Godley, Dennis, Hunter, & Bair, 2012).

Procedure exposure: Procedure exposure is an organization-level measure created for each participating organization and represents the average number of unique A-CRA procedures that the organization delivered to its respective adolescent clients. The average procedure exposure for a client across the 65 participating organizations was 9.62 ($SD=1.60$) out of a possible 19. Although there are a total of 19 A-CRA procedures, clinicians are trained to deliver procedures based on individual needs, so some clients may receive certain procedures multiple times (e.g., problem solving or communication skills) and may never receive other procedures (e.g., medication adherence and monitoring). Because each procedure involves didactic work and role playing and each session ends with a homework assignment, clinicians are usually not able to deliver more than 3 procedures in a session.

The detailed coding manual which expert raters are trained to use when rating sessions does not provide a passing score for the component labeled “Introduced procedures at appropriate times” if the clinician attempts too many procedures in one session (e.g., races through multiple procedures) (Smith, Lundy, & Gianini, 2007). As noted above, the receipt of 10 or more unique procedures during an entire treatment episode has been found to maximize client outcomes (Garner et al., 2012). Table 1 provides mean, median, minimum, and maximum percentages of clients that received each A-CRA procedure across the 65 participating organizations.

Penetration, which is similar to the concept of reach (Glasgow, 2007; Glasgow, Vogt, & Boles, 1999), is defined as the integration of a practice within a service setting and its subsystems (Proctor et al., 2011). As noted previously, penetration can be represented or conceptualized in different ways (e.g., a proportion, an absolute number). Given the information needed to compute the proportion was not available (i.e., denominator), we assessed penetration via the following two absolute measures of penetration.

Client penetration: Client penetration is an organization-level measure created for each participating organization and represents the unduplicated number of adolescent clients that the organization delivered at least one A-CRA procedure over the course of their SAMHSA/CSAT project. The average number of clients who received at least one A-CRA procedure across the 65 participating organizations was 126 ($SD=88$).

Staff penetration: Staff penetration is an organization-level measure created for each participating organization that represents the total number of “staff A-CRA certification days.” Published descriptions of the A-CRA certification process are available (Garner, Barnes, & Godley, 2009). Briefly, however, each staff delivering A-CRA to adolescent clients was required to demonstrate (via independent ratings of session audio recordings) the ability to deliver a certain number of A-CRA procedures (Godley et al., 2001) at or above the level of proficiency outlined in the A-CRA rating manual (Smith et al., 2007). Thus, for each of the staff delivering A-CRA to clients we were able to calculate the number of days in which they worked post-certification and could then sum up the total number of “staff A-CRA certification days” for each participating organization. Across the 65 participating organizations, the average number of staff A-CRA certification days was 2,340 ($SD=2,206$).

Client Outcome Measures—As noted previously, A-CRA is an EBP that has been shown to be effective in both reducing alcohol and other drug use, as well as emotional problems. Thus, given the independent importance of these two client outcomes and the extent to which adolescents present with co-morbid substance use and emotional problems, client outcomes of focus for the current study were as follows.

Substance use: Substance use is a client-reported measure that represents the number of days using any alcohol or other drugs while in the community out of the past 90 days. This measure has been shown to be consistent with both collateral reports ($kappa=.69-.92$ and agreement= $90\%-98\%$) and urine testing results ($kappa=.75-.90$ and agreement= $88\%-95\%$) (Godley, Godley, Dennis, Funk, & Passetti, 2002). Substance use, which was assessed at treatment intake and at 6 months post-intake, was examined as a client-level outcome

measure. The average number of days of substance use at intake was 28.93 ($SD=30.99$) while the average days of use at 6 months was 16.06 ($SD = 26.06$).

Emotional problems: Emotional problems is a client-reported measure that represents the proportional average of items measuring the recency (e.g., 1–3 months ago, 1–4 weeks ago, 3–7 days ago, past 2 days) and number of days (during the past 90 days) of (a) being bothered by or kept from responsibilities because of emotional problems, (b) being disturbed by memories, and (c) having problems paying attention or with self-control (Cronbach's $\alpha=.79$; seven items; ranges from 0 to 1). Cut points for severity have been empirically derived to aid clinical interpretation of this scale: low=0 to .13 (less than weekly problems); moderate=.14 to .50 (weekly problems); and high=.51 to 1.00 (daily problems). Scores greater than .13 indicate a degree of severity that warrants consideration in treatment planning. The emotional problems measure was assessed at both treatment intake and 6 months post-intake and was examined as a client-level outcome measure. The average emotional problems score at intake was 0.25 ($SD=0.20$) and at 6-month follow-up the average was 0.18 ($SD = 0.17$).

2.4 Analytic Plan

Analyses of the relationships between implementation measures and client outcomes were conducted using HLM 6 software (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2004). More specifically, a series of multilevel regression analyses (Raudenbush & Bryk, 2002) were used to regress client-level outcomes (Level 1) on organizational-level implementation measures (Level 2), with several client characteristics (e.g., age, race, gender, current criminal justice involvement) and the intake version of each client outcome measure included as controls in each respective model. Conventional $p<.05$ was used to define statistically significant relationships.

3 Results

3.1 Correlations among Implementation Measures

Table 2 shows the bivariate correlation between each of the four implementation measures examined as part of the current study. Only two correlations were statistically significant. More specifically, the two measures of fidelity (i.e., session exposure and procedure exposure) were significantly correlated with each other ($r=.56$; shared variance=31%). Additionally, the two measures of penetration (i.e., client penetration, staff penetration) were significantly correlated with each other ($r=.73$; shared variance=53%). Other correlations between measures were neither statistically significant nor meaningful (i.e., correlations were less than .10 in absolute magnitude).

3.2 Fidelity as a Predictor of Client-Level Outcomes

Table 3 presents results of the four multilevel regression analyses focused on examining the relationship between the two measures of fidelity (i.e., session exposure and procedure exposure) and the two client outcome measures of interest (i.e., substance use and emotional problems). Adjusting for client substance use at intake, as well as several client characteristics (e.g., age, race, criminal justice involvement), client substance use at follow-

up was significantly lower for treatment organizations that had higher procedure exposure ($B=-1.227$, $SE=0.583$, 95% confidence interval= -2.370 , -0.084 ; $p<.05$). Thus, each 1-unit increase in exposure (e.g., each additional added procedure) was associated with a 1.3 decrease on average in days of use over the past 3 months for these clients. This represents an effect size of $-.19$, which though relatively small, is considered to be clinically meaningful. Although in the expected direction, client substance use at follow-up was not significantly predicted by session exposure ($B=-0.116$, $SE=0.258$, $p>.05$). With regard to client emotional problems at follow-up, adjusted analyses indicated that there was no evidence of a relationship between emotional problems at follow-up and an organization's level of session exposure ($B=0.004$, $SE=0.003$) or procedure exposure ($B=0.004$, $SE=0.005$).

3.3 Penetration as a Predictor of Client-Level Outcomes

As Table 4 shows, after adjusting for the intake version of each respective client outcome measure, as well as several client characteristics (e.g., age, race, criminal justice involvement), there was no evidence of a relationship between the two examined implementation measures (i.e., client penetration, staff penetration) and the two examined outcome measures (i.e., substance use, emotional problems).

4 Discussion

Implementation outcome measures are critical for the advancement of implementation research (Proctor et al., 2011), but the availability of validated implementation measures remains limited to date. Using data from a large-scale dissemination and implementation initiative that included 65 treatment organizations, more than 300 clinicians, and approximately 6,000 adolescent clients, the current study sought to develop one or more EBIMs for one of the most widely implemented EBPs for substance use treatment (i.e., A-CRA). Multilevel analyses identified *procedure exposure* (i.e., an organization-level measure that represents the average number of unique A-CRA procedures delivered to an organization's respective clients) as one potentially meaningful EBM of implementation for A-CRA. However, none of the other three examined implementation measures (i.e., session exposure, client penetration, staff penetration) were found to be predictive of improvements in client outcomes.

Given the dearth of studies that have examined the relationship between organizational-level measures of implementation and client outcomes, the current findings contribute to A-CRA implementation research, as well as the broader implementation research field. The dearth of studies examining the relationship between organizational-level implementation measures and client outcomes makes it difficult, however, to compare and contrast the current findings with prior research. Nonetheless, the current findings do build upon research that has shown client-level A-CRA procedure exposure can predict improvements in client outcomes (Garner, 2009). For example, using a different dataset, Garner and colleagues found a client-level measure of the number of A-CRA procedures delivered to clients (i.e., the A-CRA exposure scale) mediated the relationship between treatment retention and reductions in client's days of substance use. Additionally, the current findings build upon research by Garner and colleagues (2012) that found a client-level implementation measure called Target

A-CRA (i.e., a dichotomous measure indicating whether clients received 10+ unique A-CRA procedures within 7+ treatment sessions) was significantly associated with subsequent client substance use abstinence. As such, the extent to which clients have exposure to various A-CRA treatment procedures seems to be a promising implementation measure given that it appears to be meaningful when conceptualized as a client-level implementation measure and when conceptualized as an organizational-level implementation measure.

4.1 Strengths and Limitations

A key strength of the current study is that findings are drawn from a large sample of routine substance use disorder treatment organizations, which received the same level of resources and training to implement the same EBP and collected a standardized set of implementation and outcome measures. Another key strength of the current study is that in contrast to studies that are only able to examine implementation over a relatively brief period of time (e.g., 6 months), treatment organizations were provided a relatively long period of time (i.e., 3 or more years) to implement the EBP. Finally, the study derives strength from using multilevel modeling, which has been recommended for assessing the validity of fidelity measures (Mowbray, Holter, Teague, & Bybee, 2003).

In addition to its strengths, however, the current study has a few limitations. Three key limitations of the study are that (a) client outcomes are based on client's self-report, (b) the study lacks access to other implementation outcome measures that are important to examine (e.g., acceptability, appropriateness, implementation cost), and (c) other potentially important determinants of implementation (e.g., Consolidated Framework for Implementation Research [CFIR] constructs) were not examined as part of this study. Additionally, the current study was limited to the study of a relatively basic dimension of fidelity (i.e., exposure) rather than more complex dimensions of fidelity, such as competence. That said, because this relatively basic dimension of fidelity is easier and less costly to assess than competence, examining exposure may be more useful for future research and practice. Finally, the measures of penetration were limited to those of absolute numbers (as opposed to a proportion), with the client penetration measure having a relatively low threshold (one or more A-CRA procedures).

4.2 Implications

At least two important implications can be drawn from the current research. One obvious implication is that the current research provides support for conceptual models positing an important relationship between implementation outcomes and client outcomes. A second, less obvious implication, however, is that the current research reminds us that we should not assume that implementation outcome measures are necessarily associated with subsequent client outcomes. Taken together, a key implication of the current research is that implementation researchers must carefully select implementation outcome measures that are best suited for addressing the implementation research questions of focus. For example, early-stage implementation research (e.g., Type 1 Hybrid Trials) that has a primary aim to test the effectiveness of a clinical intervention and a secondary aim to better understand the extent to which the clinical intervention (Curran, Bauer, Mittman, Pyne, & Stetler, 2012) will be acceptable to clients and/or staff would not be expected to first establish that

acceptability (i.e., an implementation outcome) is associated with subsequent client outcomes. That said, future research examining the extent to which a client's reported acceptability of the clinical intervention moderates the clinical intervention's effectiveness may be warranted. Further along, implementation research (e.g., Type 2 or Type 3 Hybrid Trials) that include a primary focus on testing strategies that lead to improvements in implementation outcomes (Curran et al., 2012) would be expected, however, to focus on improving implementation outcomes that first had been shown empirically to be related to client improvements. As an example, these findings suggest procedure exposure would most likely be related to client outcomes in a study of A-CRA.

5 Conclusions

In conclusion, the current study provides support for *procedure exposure* as an organizational-level EBM of implementation for A-CRA, which is one of the most widely disseminated and implemented EBPs for adolescent substance use disorders. As such, it seems that future efforts focused on implementing A-CRA should establish ongoing approaches to measuring and monitoring the extent to which A-CRA procedures are being delivered to clients. Procedure exposure is fortunately one of the more pragmatic implementation outcomes (see Lewis et al., 2015 for more information on pragmatic implementation measures) because it is something that can (and should be) documented by the implementation staff. Nonetheless, we encourage future efforts to seek ways to measure and monitor other potentially key implementation measures (e.g., competence, proportional measures of penetration), which may be even better predictors of subsequent client outcomes.

As noted by Durlak and DuPre (2008), "science cannot study what it cannot measure accurately and cannot measure what it does not define." Given the field of implementation research has now begun to reach consensus on defining the key conceptual models and measurement constructs (Damschroder et al., 2009; Proctor et al., 2011; 2009), there is now a critical need to increase efforts to develop EBMI (i.e., measures that are reliable and valid predictors of important client outcomes). Such efforts will be challenging, of course, but because reliable and valid measures are central to good science, developing EBMI must become a top priority for future implementation research. Although research will likely need to develop EBMI for a specific EBPs, as the literature on EBMI grows, meta-analytic techniques may enable researchers to explore which EBMI are generally more predictive of client outcomes and/or are generalizable across different types of EBPs. Our hope is that the current research will serve as a prototype for future research in this area (i.e., the examination of the relationship between implementation outcomes and improvements in key client outcomes), which in turn will help implementation research achieve one of its ultimate goals: identifying effective and cost-effective strategies to improve EBP implementation and sustainability within routine practice settings.

Acknowledgments

This work was supported by the National Institute on Alcohol Abuse and Alcoholism (NIAAA; R01-AA021217) and the Substance Abuse and Mental Health Services Administration's Center for Substance Abuse Treatment

(SAMHSA/CSAT; contract no. 270-07-0191). The content is solely the responsibility of the authors and does not necessarily represent the official views of the government.

References

- Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: Combining elements of clinical effectiveness and implementation research to enhance public health impact. *Med Care*. 2012; 50(3):217–226. DOI: 10.1097/MLR.0b013e3182408812 [PubMed: 22310560]
- Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implement Sci*. 2009; 4:50.doi: 10.1186/1748-5908-4-50 [PubMed: 19664226]
- Dane AV, Schneider BH. Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clin Psychol Rev*. 1998; 18(1):23–45. [PubMed: 9455622]
- Dennis M, Godley SH, Diamond G, Tims FM, Babor T, Donaldson J, Liddle H, Titus JC, Kaminer Y, Webb C, Hamilton N, Funk R. The Cannabis Youth Treatment (CYT) Study: Main findings from two randomized trials. *J Subst Abuse Treat*. 2004; 27(3):197–213. DOI: 10.1016/j.jsat.2003.09.005 [PubMed: 15501373]
- Dennis ML, Funk R, Godley SH, Godley MD, Waldron H. Cross validation of the alcohol and cannabis use measures in the Global Appraisal of Individual Needs (GAIN) and Timeline Followback (TLFB; Form 90) among adolescents in substance abuse treatment. *Addiction*. 2004; 99(Suppl 2):120–128. [PubMed: 15488110]
- Dennis, ML.; Titus, JC.; White, M.; Unsicker, J.; Hodgkins, D. Global Appraisal of Individual Needs (GAIN): Administration guide for the GAIN and related measures. Version 5. Chestnut Health Systems; Bloomington, IN: 2003.
- Durlak JA, DuPre EP. Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am J Community Psychol*. 2008; 41(3-4):327–350. DOI: 10.1007/s10464-008-9165-0 [PubMed: 18322790]
- Eccles MP, Mittman BS. Welcome to implementation science. *Implement Sci*. 2006; 1(1):1–3. DOI: 10.1186/1748-5908-1-1
- Elbardissi AW, Duclos A, Rawn JD, Orgill DP, Carty MJ. Cumulative team experience matters more than individual surgeon experience in cardiac surgery. *J Thorac Cardiovasc Surg*. 2013; 145(2):328–333. DOI: 10.1016/j.jtcvs.2012.09.022 [PubMed: 23083794]
- Garner BR. Research on the diffusion of evidence-based treatments within substance abuse treatment: A systematic review. *J Subst Abuse Treat*. 2009; 36(4):376–399. DOI: 10.1016/j.jsat.2008.08.004 [PubMed: 19008068]
- Garner BR, Barnes B, Godley SH. Monitoring fidelity in the Adolescent Community Reinforcement Approach (A-CRA): The training process for A-CRA raters. *J Behav Anal Health Sports Fit Med*. 2009; 2(1):43–54. [PubMed: 21533009]
- Garner BR, Godley MD, Funk RR, Dennis ML, Godley SH. The impact of continuing care adherence on environmental risks, substance use, and substance-related problems following adolescent residential treatment. *Psychol Addict Behav*. 2007; 21(4):488–497. DOI: 10.1037/0893-164x.21.4.488 [PubMed: 18072831]
- Garner BR, Godley SH, Dennis ML, Hunter B, Bair C, Godley MD. Using pay for performance to improve treatment implementation for adolescent substance use disorders: Results from a cluster randomized trial. *Arch Pediatr Adolesc Med*. 2012; 166:938–944. DOI: 10.1007/archpediatrics.2012.802 [PubMed: 22893231]
- Garner BR, Godley SH, Funk RR, Dennis ML, Smith JE, Godley MD. Exposure to Adolescent Community Reinforcement Approach treatment procedures as a mediator of the relationship between adolescent substance abuse treatment retention and outcome. *J Subst Abuse Treat*. 2009; 36(3):252–264. DOI: 10.1016/j.jsat.2008.06.007 [PubMed: 18715742]
- Glasgow RE. eHealth evaluation and dissemination research. *Am J Prev Med*. 2007; 32(5 Suppl):S119–126. DOI: 10.1016/j.amepre.2007.01.023 [PubMed: 17466816]

- Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: The RE-AIM framework. *Am J Public Health*. 1999; 89(9):1322–1327. [PubMed: 10474547]
- Godley MD, Godley SH, Dennis ML, Funk R, Passetti LL. Preliminary outcomes from the assertive continuing care experiment for adolescents discharged from residential treatment. *J Subst Abuse Treat*. 2002; 23(1):21–32. [PubMed: 12127465]
- Godley MD, Godley SH, Dennis ML, Funk RR, Passetti LL. The effect of assertive continuing care on continuing care linkage, adherence and abstinence following residential treatment for adolescents with substance use disorders. *Addiction*. 2007; 102(1):81–93. DOI: 10.1111/j.1360-0443.2006.01648.x [PubMed: 17207126]
- Godley SH, Garner BR, Smith JE, Meyers RJ, Godley MD. A large-scale dissemination and implementation model for evidence-based treatment and continuing care. *Clin Psychol (New York)*. 2011; 18(1):67–83. DOI: 10.1111/j.1468-2850.2011.01236.x [PubMed: 21547241]
- Godley SH, Hunter BD, Fernandez-Artamendi S, Smith JE, Meyers RJ, Godley MD. A comparison of treatment outcomes for adolescent community reinforcement approach participants with and without co-occurring problems. *J Subst Abuse Treat*. 2014; 46(4):463–471. DOI: 10.1016/j.jsat.2013.10.013 [PubMed: 24462478]
- Godley, SH.; Meyers, RJ.; Smith, JE.; Godley, MD.; Titus, JC.; Karvinen, T.; Dent, G.; Passetti, LL.; Kelberg, P. Cannabis youth treatment series volume 4: The adolescent community reinforcement approach (DHHS Publication No. SMA 01-3489). Center for Substance Abuse Treatment, Substance Abuse and Mental Health Services Administration; Rockville, MD: 2001.
- Hunter SB, Ayer L, Han B, Garner BR, Godley SH. Examining the sustainment of the Adolescent-Community Reinforcement Approach in community addiction treatment settings: protocol for a longitudinal mixed method study. *Implement Sci*. 2014; 9:104. [PubMed: 25116509]
- Institute of Medicine. Bridging the gap between practice and research: forging partnerships with community-based drug and alcohol treatment. National Academy Press; Washington, DC: 1998.
- Institute of Medicine. Crossing the quality chasm: A new health system for the 21st century. National Academy Press; Washington, DC: 2001.
- Keith RE, Hopp FP, Subramanian U, Wiitala W, Lowery JC. Fidelity of implementation: Development and testing of a measure. *Implement Sci*. 2010; 5:99.doi: 10.1186/1748-5908-5-99 [PubMed: 21192817]
- Kerner JF. Knowledge translation versus knowledge integration: A "funder's" perspective. *J Contin Educ Health Prof*. 2006; 26(1):72–80. DOI: 10.1002/chp.53 [PubMed: 16557513]
- Klein KJ, Sorra JS. The challenge of innovation implementation. *Academy of Management Review*. 1996; 21(4):1055–1080. DOI: 10.2307/259164
- Lewis CC, Weiner BJ, Stanick C, Fisher SM. Advancing implementation science through measure development and evaluation: A study protocol. *Implement Sci*. 2015; 10(102):1–10. DOI: 10.1186/s13012-015-0287-0 [PubMed: 25567289]
- Martinez RG, Lewis CC, Weiner BJ. Instrumentation issues in implementation science. *Implement Sci*. 2014; 9:118.doi: 10.1186/s13012-014-0118-8 [PubMed: 25185799]
- Mesman R, Westert GP, Berden BJ, Faber MJ. Why do high-volume hospitals achieve better outcomes? A systematic review about intermediate factors in volume-outcome relationships. *Health Policy*. 2015; 119(8):1055–1067. DOI: 10.1016/j.healthpol.2015.04.005 [PubMed: 25958187]
- Miller WR, Yahne CE, Moyers TB, Martinez J, Pirritano M. A randomized trial of methods to help clinicians learn motivational interviewing. *J Consult Clin Psychol*. 2004; 72(6):1050–1062. DOI: 10.1037/0022-006x.72.6.1050 [PubMed: 15612851]
- Mitchell SA, Fisher CA, Hastings CE, Silverman LB, Wallen GR. A thematic analysis of theoretical models for translational science in nursing: mapping the field. *Nurs Outlook*. 2010; 58(6):287–300. DOI: 10.1016/j.outlook.2010.07.001 [PubMed: 21074646]
- Mowbray CT, Holter MC, Teague GB, Bybee D. Fidelity criteria: Developmental, measurement, and validation. *Am J Eval*. 2003; 24(3):315–340.
- Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, Griffey R, Hensley M. Outcomes for implementation research: Conceptual distinctions, measurement challenges, and

research agenda. *Adm Policy Ment Health*. 2011; 38(2):65–76. DOI: 10.1007/s10488-010-0319-7 [PubMed: 20957426]

Proctor EK, Landsverk J, Aarons G, Chambers D, Glisson C, Mittman B. Implementation research in mental health services: An emerging science with conceptual, methodological, and training challenges. *Adm Policy Ment Health*. 2009; 36(1):24–34. DOI: 10.1007/s10488-008-0197-4 [PubMed: 19104929]

Proctor, EK.; Powell, BJ.; Feely, M. Measurement in dissemination and implementation science. In: Beidas, RS.; Kendall, PC., editors. *Dissemination and implementation of evidence-based practices in child and adolescent mental health*. Oxford University Press; New York, NY: 2014. p. 22-43.

Raudenbush, SW.; Bryk, AS. *Hierarchical linear models*. Sage; Thousand Oaks, CA: 2002.

Raudenbush, SW.; Bryk, AS.; Cheong, YF.; Congdon, R.; du Toit, M. *HLM 6: Hierarchical linear and nonlinear modeling*. Scientific Software International, Inc; Lincolnwood, IL: 2004.

Sholomskas DE, Syracuse-Siewert G, Rounsaville BJ, Ball SA, Nuro KF, Carroll KM. We don't train in vain: A dissemination trial of three strategies of training clinicians in cognitive-behavioral therapy. *J Consult Clin Psychol*. 2005; 73(1):106–115. DOI: 10.1037/0022-006X.73.1.106 [PubMed: 15709837]

Simpson DD. A conceptual framework for transferring research to practice. *J Subst Abuse Treat*. 2002; 22(4):171–182. [PubMed: 12072162]

Smith, JE.; Lundy, SL.; Gianini, L. *Community reinforcement approach (CRA) and adolescent community reinforcement approach (A-CRA) coding manual*. University of New Mexico; Albuquerque, NM: 2007.

Tabak RG, Khoong EC, Chambers DA, Brownson RC. Bridging research and practice: models for dissemination and implementation research. *Am J Prev Med*. 2012; 43(3):337–350. DOI: 10.1016/j.amepre.2012.05.024 [PubMed: 22898128]

Wilson PM, Petticrew M, Calnan MW, Nazareth I. Disseminating research findings: what should researchers do? A systematic scoping review of conceptual frameworks. *Implement Sci*. 2010; 5(91):1–16. DOI: 10.1186/1748-5908-5-91 [PubMed: 20047652]

Highlights

- Sought to develop evidence-based measures of implementation (EBMIs) for A-CRA.
- Dataset included 65 treatment organizations, 308 clinicians, and 5,873 clients.
- Procedure exposure was found to be an EBM of implementation for A-CRA.
- Study may serve as prototype for research identifying EBMs of implementation.

Table 1

Client receipt of A-CRA procedure across the 65 participating organizations

Procedure	Mean	Median	Minimum	Maximum
A-CRA overview	84%	89%	42%	100%
Happiness scale	85%	87%	55%	100%
Goals of counseling	82%	84%	39%	100%
Functional analysis of substance abuse	73%	73%	35%	97%
Problem-solving skills	68%	68%	34%	97%
Homework review	68%	90%	0%	100%
Communication skills	66%	67%	27%	94%
Pro-social recreation	54%	55%	19%	84%
Functional analysis of pro-social behavior	53%	54%	17%	93%
Caregiver overview	47%	49%	7%	87%
Relapse prevention skills	46%	43%	17%	86%
Drink and drug refusal skills	45%	46%	7%	77%
Caregiver relationship skills	38%	39%	6%	79%
Systematic encouragement	37%	34%	1%	72%
Job-seeking skills	37%	37%	6%	72%
Anger management skills	36%	33%	0%	82%
Sobriety sampling	29%	25%	1%	68%
Medication compliance	4%	1%	0%	23%
Couples relationship skills	2%	1%	0%	41%

Table 2

Correlations among implementation measures

Implementation measure	Session exposure	Procedure exposure	Client penetration	Staff penetration
Session exposure	–	–	–	–
Procedure exposure	.56	–	–	–
Client penetration	.09	–.01	–	–
Staff penetration	.07	.04	.73	–

Note: Bold indicates $p < .05$.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Fidelity as a predictor of client-level outcomes

	Days using (n=5,685)		Emotional problems (n=5,697)	
	<i>B</i> (<i>SE</i>)	95% CI lower, upper	<i>B</i> (<i>SE</i>)	95% CI lower, upper
Session exposure model				
<i>Organizational-level measure</i>				
Session exposure	-0.116 (0.258)	(-0.622, 0.390)	0.004 (0.003)	(-0.001, 0.010)
<i>Client-level measures</i>				
Outcome measure at intake	0.227 (0.013)	(0.201, 0.252)	0.323 (0.014)	(0.295, 0.351)
Age	0.364 (0.176)	(0.018, 0.709)	-0.002 (0.002)	(-0.005, 0.002)
African American	0.558 (1.064)	(-1.528, 2.645)	-0.012 (0.006)	(-0.025, -0.0001)
Hispanic	0.022 (0.943)	(-1.826, 1.871)	-0.024 (0.007)	(-0.038, -0.010)
Other race	-0.644 (0.925)	(-2.456, 1.168)	-0.006 (0.006)	(-0.018, 0.006)
Female	-2.136 (0.846)	(-3.793, -0.479)	0.025 (0.007)	(0.012, 0.038)
Criminal justice involvement	-0.590 (0.711)	(-1.984, 0.805)	0.006 (0.005)	(-0.003, 0.015)
Procedure exposure model				
<i>Organizational-level measure</i>				
Procedure exposure	-1.227 (0.583)	(-2.370, -0.084)	0.004 (0.005)	(-0.006, 0.013)
<i>Client-level measures</i>				
Outcome measure at intake	0.227 (0.013)	(0.201, 0.252)	0.323 (0.014)	(0.295, 0.351)
Age	0.366 (0.176)	(0.021, 0.711)	-0.002 (0.002)	(-0.005, 0.002)
African American	0.466 (1.067)	(-1.623, 2.556)	-0.013 (0.006)	(-0.025, -0.0001)
Hispanic	0.146 (0.944)	(-1.704, 1.996)	-0.024 (0.007)	(-0.039, -0.010)
Other race	-0.660 (0.929)	(-2.482, 1.160)	-0.006 (0.006)	(-0.018, 0.006)
Female	-2.170 (0.845)	(-3.827, -0.514)	0.025 (0.007)	(0.012, 0.038)
Criminal justice involvement	-0.555 (0.714)	(-1.955, 0.845)	0.006 (0.005)	(-0.004, 0.015)

Note: **Bold** indicates $p < .05$. Caucasian was the referent group for race.

Table 4

Penetration as a predictor of client-level outcomes

	Days using (n=5,685)		Emotional problems (n=5,697)	
	<i>B</i> (<i>SE</i>)	95% CI lower, upper	<i>B</i> (<i>SE</i>)	95% CI lower, upper
Client penetration model				
<i>Organizational-level measure</i>				
Client penetration	0.003 (0.007)	(-0.011, 0.016)	-0.00004	(-0.0001, 0.0001)
<i>Client-level measures</i>				
Outcome measure at intake	0.227 (0.013)	(0.201, 0.252)	0.323 (0.014)	(0.295, 0.351)
Age	0.363 (0.176)	(0.018, 0.709)	-0.002 (0.002)	(-0.005, 0.002)
African American	0.574 (1.064)	(-1.511, 2.659)	-0.013 (0.006)	(-0.025, -0.0002)
Hispanic	0.026 (0.948)	(-1.833, 1.884)	-0.024 (0.007)	(-0.039, -0.010)
Other race	-0.637 (0.927)	(-2.454, 1.180)	-0.006 (0.006)	(-0.018, 0.006)
Female	-2.140 (0.842)	(-3.791, -0.489)	0.025 (0.007)	(0.012, 0.038)
Criminal justice involvement	-0.585 (0.708)	(-1.972, 0.802)	0.006 (0.005)	(-0.004, 0.015)
Staff penetration model				
<i>Organizational-level measure</i>				
Staff penetration	0.0001 (0.0002)	(-0.0004, 0.0005)	0.0000 (0.00001)	(-0.00001, 0.00001)
<i>Client-level measures</i>				
Outcome measure at intake	0.227 (0.013)	(0.201, 0.252)	0.323 (0.014)	(0.295, 0.351)
Age	0.363 (0.176)	(0.018, 0.709)	-0.002 (0.002)	(-0.005, 0.002)
African American	0.575 (1.064)	(-1.510, 2.662)	-0.013 (0.006)	(-0.025, -0.0002)
Hispanic	0.032 (0.947)	(-1.824, 1.889)	-0.024 (0.007)	(-0.039, -0.010)
Other race	-0.636 (0.927)	(-2.454, 1.181)	-0.006 (0.006)	(-0.018, 0.006)
Female	-2.141 (0.843)	(-3.793, -0.488)	0.025 (0.007)	(0.012, 0.038)
Criminal justice involvement	-0.583 (0.709)	(-1.972, 0.806)	0.006 (0.005)	(-0.004, 0.015)

Note: **Bold** indicates $p < .05$. Caucasian was the referent group for race.