

ORIGINAL RESEARCH

OPEN ACCESS
Full open access to this and thousands of other papers at <http://www.la-press.com>.

Development and Initial Validation of a Client-Rated MET-CBT Adherence Measure

Wendy R. Ulaszek^{1,2}, Hsiu-Ju Lin^{1,2}, Linda K. Frisman^{1,2}, Susan Sampl³, Susan Harrington Godley⁴, Karen L. Steinberg-Gallucci³, Jody L. Kamon⁵ and Margaret O'Hagan-Lynch²

¹University of Connecticut School of Social Work, ²Connecticut Department of Mental Health and Addiction Services Research Division, ³University of Connecticut Health Center, ⁴Chestnut Health Systems, ⁵University of Vermont. Corresponding author email: wendy.ulaszek@uconn.edu

Abstract: Traditional mechanisms for rating adherence or fidelity are labor-intensive. We developed and validated a tool to rate adherence to Motivational Enhancement Therapy—Cognitive Behavioral Treatment (MET-CBT) through anonymous client surveys. The instrument was used to survey clients in 3 methadone programs over 2 waves. Explanatory and Confirmatory Factor Analyses were used to establish construct validity for both MET and CBT. Internal consistency based on Cronbach's alpha was within adequate range ($\alpha > 0.70$) for all but 2 of the subscales in one of the samples. Consensus between clients' ratings ($r_{wg(j)}$ scores) were in the range of 0.6 and higher, indicating a moderate to strong degree of agreement among clients' ratings of the same counselor. These results suggest that client surveys could be used to measure adherence to MET-CBT for quality monitoring that is more objective than counselor self-report and less resource-intensive than supervisor review of taped sessions. However, additional work is needed to develop this scale.

Keywords: MET-CBT, fidelity, implementation, client ratings

Substance Abuse: Research and Treatment 2012:6 85–94

doi: [10.4137/SART.S9896](https://doi.org/10.4137/SART.S9896)

This article is available from <http://www.la-press.com>.

© the author(s), publisher and licensee Libertas Academica Ltd.

This is an open access article. Unrestricted non-commercial use is permitted provided the original work is properly cited.



Introduction

The gap between science and standard practice is especially wide in the area of substance abuse treatment.¹⁻³ More and more states are mandating that behavioral health care providers, like other health care providers, use evidence-based practices (EBPs).^{1,4} However, EBPs that are introduced to practitioners without ensuring correct implementation are limited in their usefulness and effectiveness.⁵⁻⁷ Therapist manuals and one-time workshops by themselves are known to be ineffective in helping practitioners utilize new skills.^{1,8-10} Even when new EBPs are implemented successfully, adherence drift can become a serious problem.¹ Fidelity monitoring and feedback can help curb this type of drift, and is frequently used in the context of research studies.

Several models have been described to monitor use of EBPs. The “gold standard” of fidelity monitoring is expert or supervisor review of video- or audio-taped therapy sessions and/or live observation.¹¹ In the real world, it is not feasible for many agencies to use these costly and time-consuming measures to monitor quality of treatment,¹² they are more likely to rely on counselors’ self-reported behavior. However, counselors’ self-reported fidelity to new therapies does not correlate well to actual proficiency or use of the skills.¹³⁻¹⁵ Client ratings may better fit the need for low-burden assessments of EBP adherence, but there is insufficient evidence of their validity and reliability.¹⁶

In this paper, we describe an innovative fidelity monitoring technique: the use of an anonymous, self-administered form to be completed by clients, and the development and initial validation of this fidelity adherence measure. In particular, we developed an instrument for clients to rate their counselors’ use of Motivational Enhancement Therapy and Cognitive Behavioral Therapy (MET-CBT). Both MET and CBT are empirically validated addiction treatment methods,^{4,11,17,18} but counselors’ adherence to these models is difficult to measure. Indeed, randomized controlled studies conducted to show the efficacy of motivational techniques have often lacked measurement of counselors’ use of the skills involved.¹⁹⁻²¹ The purpose of this article is to describe the psychometric properties of a client-rated fidelity measure of MET-CBT.

The implementation study context

The development of the MET-CBT client-rated fidelity adherence measure was part of a project testing a model of EBP dissemination. The study was funded by the National Institute on Drug Abuse (NIDA) and conducted from August 1, 2005 to July 29, 2008, in three methadone clinics which are all part of a single large urban addiction treatment agency. The study protocol was approved by the Institutional Review Board (IRB) of the Connecticut Department of Mental Health and Addiction Services, where the first three authors serve in the Research Division. The featured treatment was a blended model of (1) MET, which is especially useful in improving clients’ engagement and motivation to change their substance use behaviors,²² and (2) CBT, which gives clients the needed skills to carry out these changes and to address problems that lead to substance abuse.²³ The project employed on-site training of all counseling staff and supervisors. Supervisors received additional training in how to reinforce MET-CBT skills. In addition several accepted implementation techniques were applied. These included attending to barriers to organizational change; involving all levels of staff in the change process; adaptation of the EBPs to the local setting’s procedures; and training supervisors to provide regular feedback to counselors on adherence and skillfulness. Moreover, a member of the research team, termed an ‘implementation shepherd’ was identified to facilitate monthly advisory agency staff meetings during which barriers to implementation were resolved.

Method

Development of the MET-CBT client-rated adherence measure

Initial generation of items for the client-rated fidelity measure was based on review of the MET-CBT literature and clinical experience of two of the authors, both clinical psychologists, and one of whom is an expert trainer in MET-CBT. We drafted client statements that might reflect the principles of MET (eg, employing empathy; rolling with resistance) and CBT (eg, coping with risk; developing refusal skills) from a client’s perspective. Items were written without clinical jargon and were meant to reflect the behaviors and attitudes of



clinicians who are using MET-CBT skills in counseling sessions. Initially, enough items in each domain were included so that analyses could identify the strongest items for inclusion in a future, shorter version.

The first draft of items was sent to four national MET-CBT experts. Experts were asked to independently rate each MET-CBT client fidelity item on a scale of 1–5. Scale values were defined as: 1 = Definitely omit (neither item nor its reversal correctly describes MET or CBT principles); 2 = Probably should omit (unnecessary, confusing, and/or does not discriminate MET or CBT from other techniques); 3 = Keep but re-word (discriminates MET or CBT fairly well, but not clearly); 4 = Probably can keep (a good item, as it discriminates MET-CBT adequately, fairly clear, but may need minor re-wording); 5 = Keep as written (Very good item, highly related to MET-CBT, discriminates well, and clearly worded). We also asked each expert to nominate 15 items to be omitted and to suggest wording changes that would help to make items clearer. We specifically targeted a 6th grade reading level. We calculated the discrepancy between expert ratings and examined 12 items with averages less than or equal to 3.0. Also, based on theory and expert opinion, we established 6 subscales that seemed to capture the essence of using MET; they were: (1) Client Centered Focus, (2) Strengths Based, Self-Efficacy, (3) Empathy and Acceptance, (4) Avoiding Argumentation, (5) Rolling with Resistance, and (6) Developing Discrepancy. Likewise, we established two theory-driven subscales for CBT: (1) Functional Analysis and (2) Skills Training.

We presented the draft client-rated adherence measure to agency advisory board members. Per advisory board recommendation, we administered a pilot of the survey to 14 group therapy patients at the agency. We then reviewed the pilot responses and revised the measure per client and advisory board members' feedback. Changes included adding a "Not Applicable/Don't Know" response and making the lay-out more user-friendly. Finally, we obtained approval from the IRB for the anonymous client survey, the length of which was two double-sided pages.

Data collection procedures

Prior to the staff training in MET-CBT, clients of the affected methadone programs were asked to complete

the MET-CBT client-rated adherence scale. We refer to this phase as Wave 1 data collection. Patients eligible for responding to the anonymous survey were those patients in treatment for at least two months and who were seen by an agency counselor at least monthly. Respondents had to be at least 18 years old, able to read in English, and have no conservator. Confidential lists of eligible patients were generated internally by the agency clinical coordinator and distributed to the three program sites. Receptionists at each of the three sites were given a standard protocol and trained on how to ask identified patients to complete the survey and to cross off names from the confidential eligibility lists when the person either completed a survey, or refused to complete a survey, or withdrew from treatment. Moreover, administrative agency receptionists were trained to place the date and the name of the client's primary counselor on the survey before handing it to the client to complete. People who had difficulty reading or seeing the survey were invited to meet with a research assistant for phone or interview administration. The study participants were asked to insert the completed surveys into a locked box that could only be opened by research staff. The receptionists gave clients who completed the surveys a \$10 gift certificate to a local store as an incentive for participation. Given the anonymity of the survey and the careful methods of retrieval of surveys, a waiver of informed consent was granted by the IRB.

After the initial MET-CBT training was completed, the data collection procedure was repeated, using the same eligibility criteria for clients as in Wave 1. No effort was made to specifically follow up with Wave 1 respondents or to exclude them. Thus, these samples should be considered separate cross-sections of agency clients. Approximately 6 months later, the Wave 3 data collection was initiated, following the same eligibility criteria and procedures used in Wave 1 and Wave 2.

Participants

A total of 610 participants completed the client-rated MET-CBT adherence measure. Two were removed from the data analysis due to response bias (circling the same response throughout the survey). Another four were removed due to incomplete surveys, where more than two-thirds of the survey items were missing.



Thus, a total of 604 clients were included in the final analyses (194 for Wave 1, 205 for Wave 2, and 205 for Wave 3).

Results

Construct validity

Exploratory Factor Analyses (EFAs). A multi-step approach was adopted to evaluate the construct validity for the client level CBT and MET adherence scales. First, we determined that there were no significant differences between the background information for clients in Wave 1 and Wave 2, and the total scale score did not change. Therefore, to increase the sample size for the exploratory factor analyses (EFA), we combined these two waves of data. Table 1 shows the background information by sample groups. EFA was then employed to evaluate the factor structure using the combined Wave 1 and Wave 2 data (now termed the development sample). The factor solutions from the EFA were verified through Confirmatory Factor Analyses (CFA) using the Wave 3 data (the cross-validation sample).

Since the goal for the EFA is to identify latent variables, principal-axis factor analysis (PAF), also known as common factor analysis, was selected as the factor extraction method.²⁴ Because we expected

the factors to be correlated, we used oblique rotation rather than orthogonal. The SPSS 15.0 procedure FACTOR (SPSS, 2006) was used to perform the EFA. To determine the number of factors to retain, we used the following criteria: (1) eigen values > 1.0; (2) last substantial drop in the scree plot; (3) interpretability of the solution; and (4) minimum of three items per factor.

The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity indicated that both the MET and CBT scales were psychometrically adequate for factor analysis (for CBT, KMO = 0.94, Bartlett's $\chi^2 = 3276.20$, $df = 91$, $P < 0.000$; for MET, KMO = 0.94, Bartlett's $\chi^2 = 8263.84$, $df = 703$, $P < 0.000$).

For CBT, the initial factor analysis yielded a three-factor solution; however, three items had unclear factor loadings. For example, one item did not load onto any factor, and two items had high factor loadings for more than one factor, which implied that the factors were not distinct for these items. As suggested by Kahn,²⁴ these items were removed to obtain a clearer factor solution. The factor analysis was re-run to reveal a two-factor solution. The first factor accounted for 55.25% of the total variance, and the second factor added an additional 7.91% variance.

Table 1. Background information for the development and cross-validation samples.

	Development sample		Cross-validation sample		
	N	% SD	N	% SD	
Age					$\chi^2(4) = 18.67, P = 0.001$
Under 20	2	0.5%	7	3.4%	
21–30	102	25.6%	71	34.6%	
31–40	127	31.8%	65	31.7%	
40–50	142	35.6%	47	22.9%	
50+	26	6.5%	15	7.3%	
Female*	No data	No data	96	46.8%	–
Race					ns
White	294	74.1%	153	74.6%	
Hispanic	39	9.8%	26	12.7%	
Black	51	12.8%	18	8.8%	
Other	13	3.3%	8	3.9%	
Time worked with the counselor (in months)	5.52	7.25	9.95	19.99	$t(593) = -3.02, P < 0.000$
# of individual sessions per month	2.07	2.24	2.34	1.36	ns
# of group sessions per month	1.96	1.96	1.84	2.09	ns

Note: *Gender was not measured in the development sample.



Items that loaded onto the first factor reflected the combined concept of “Functional Analysis/Coping with Risk” (eg, “My counselor helps me figure out other things to do with my time instead of using”), and items loading onto the second factor represented the “Developing Life Skills” domain (eg, “In my sessions, I learn how to solve problems by breaking them down into steps”).

The EFA for the MET scale revealed a potential five factor solution. The five-factor solution accounted for 56.29% of the variance, and reflected the following five domains: “Support Self-Efficacy/Elicit Change Talk” (eg, “My counselor and I talk about what I want for my future”); “Avoid Argumentation” (eg, “I feel like I have to defend myself to my counselor [reverse scored]”); “Roll with Resistance” (eg, “My counselor gets mad if I don’t follow our treatment plan” [reverse scored]); “Client Centered Perspective” (eg, “I don’t really think that my counselor understands me” [reverse scored]); and “Express Empathy/Acceptance” (eg, “My counselor tries to see things from my point of view”).

Confirmatory Factor Analyses (CFA). To continue examining the construct validity for the MET-CBT subscales, we performed Confirmatory Factor Analyses (CFAs) on the cross-validation sample. These CFAs allowed us to compare the model fit for the factor structures derived from the EFAs and factors derived from a priori theory-based and expert opinion. CFAs models were estimated using the AMOS 4.0 program²⁵ with a full information maximum likelihood (FIML) solution. The model fit was evaluated using various fit indices. The non-significant chi-square indicated a good overall model fit. Also, we applied the Hu et al²⁶ standards, which suggest accepting a Tucker-Lewis Index (TLI, also known as a non-normed fit index: NNFI) >0.95, and a comparative fit index (CFI) >0.95. Applying the Steiger²⁷ recommendations, we checked for a root mean square error (RMSEA) of <0.060. With respect to the latter statistic, most researchers consider a RMSEA less than 0.08 as an acceptable model fit, and a value less than 0.06 a very good model fit.²⁸ Since not all the models are nested models, we also examined the Akaike information criterion (AIC), for which lower scores indicate a better fit when comparing two non-nested models.²⁹ Factors were free to correlate with each other, but each item was constrained to load on

only one factor. Factor variances were fixed at 1.0 and factor loadings were not constrained.

The CFA results suggest that the two-factor model derived from the EFA (χ^2 (76) = 183.42, P = 0.00, TLI = 0.98, CFI = 0.99, RMSEA = 0.083, and AIC = 269.42) fit the data better than either the one factor model (χ^2 (77) = 190.22, P = 0.00, TLI = 0.98, CFI = 0.99, RMSEA = 0.085, and AIC = 269.42; $\Delta\chi^2$ (1) = 6.8, P < 0.01), or the two-factor model based on the theory (χ^2 (103) = 498.35, P = 0.00, TLI = 0.98, CFI = 0.99, RMSEA = 0.098, and AIC = 596.35). The two-factor model based on the EFA also has the smallest AIC value compared to the other models. Although the RMSEA for the two-factor EFA model is slightly over 0.08, it has the smallest RMSEA value compared to the two alternative models. In sum, the results from the CFA reveals that the two-factor model based on the EFA has the best model fit indices.

For the MET scale, model fit indices for the five-factor model derived from the EFA showed the best fit for the data (χ^2 (517) = 912.80, P = 0.00, TLI = 0.98, CFI = 0.98, RMSEA = 0.061, and AIC = 1136.80), compared to either the one-factor model (χ^2 (527) = 1390.02, P = 0.00, TLI = 0.96, CFI = 0.96, RMSEA = 0.090, and AIC = 1594.02), or the six-factor model based on the theory (χ^2 (514) = 1358.72, P = 0.00, TLI = 0.96, CFI = 0.96, RMSEA = 0.090, and AIC = 1588.72). The TLI and CFI for the five-factor EFA model are close to 1, and the RMSEA value is smaller than 0.08 (0.063). These fit indices suggest that the five-factor model is preferred.

Tables 2A and B show the CFA factor loadings for the two-factor model of CBT, and the five-factor model of MET, both of which were derived from the EFA. All factor loadings were statistically significant at the P = 0.05 level. With the exception of the MET item “My counselor pushes me to change my life,” which had a low factor loading of 0.12, the remaining factor loadings were satisfactory. For CBT, they ranged from 0.41 to 0.76, with a mean of 0.59. For MET, they ranged from 0.28 to 0.79, with a mean of 0.60. We found that removing the item “My counselor pushes me to change my life” from the CFAs would slightly improve the factor model fit for the MET 5-factor model (χ^2 = 837.71, df = 485, RMSEA = 0.06, CFI = 0.98, TLI = 0.98, AIC = 1055.71). However, item analyses also showed that the reliability coefficient for the Role with Resistance subscale would not

**Table 2A.** Confirmatory Factor Analysis of the CBT client-rated fidelity scale.

Item	Loading
Factor 1: Functional Analysis/Coping with Risk (FA/CR)	
My counselor helps me figure out other things to do with my time instead of using.	0.76
My counselor helps me plan for risky situations in the future.	0.70
My counselor and I talk about how I deal with people, places, and things that put me at risk for using.	0.67
My counselor and I talk about ways I can avoid situations that make me want to use.	0.64
My counselor tells me to try new things, or get new hobbies.	0.64
My counselor and I talk about what happens in my life when I use drugs and/or alcohol.	0.62
My counselor helps me to talk about my thoughts and feelings when I use.	0.59
My counselor helps me to look at thoughts and feelings that go with wanting to use.	0.54
My counselor helps me figure out which people, places and things put me at risk for using.	0.47
Factor 2: Developing Life Skills (DLS)	
In the sessions, I learn how to solve problems by breaking them down into steps.	0.72
During my therapy sessions, I practice how to turn down drugs or alcohol.	0.55
My counselor helps me talk about other issues, like where to live.	0.50
My counselor and I talk about problems with work, or finding a job.	0.45
My counselor and I use role-playing to practice new skills.	0.41

have been improved by removing this item from the scale. Thus, we retained this item for our analyses.

Means and standard deviations for the final CBT and MET subscales are presented in Table 3. Tables 4A and B show the inter-correlations among the CBT and MET subscales, as well as the total scores for the development and cross-validation samples.

Internal reliability

Internal consistencies were examined using Cronbach's alpha for the total scores and subscales, based on the models that had the best model fit statistics from the CFAs. Table 4 presents the alphas for the developmental and cross-validation samples. For the developmental sample, Cronbach's alphas were all within an adequate range ($\alpha > 0.70$).³⁰ In contrast, two subscales for the cross-validation sample had low alphas: The Roll with Resistance subscale was 0.37, and the Acceptance subscale was 0.57. The large discrepancy of alpha values for the Roll with Resistance subscale and the Acceptance subscale between the development sample (α : RR = 0.72, and AC = 0.71), and cross-validation samples (α : RR = 0.37, and AC = 0.57) could be the result of sampling variation. To test this explanation, we compared characteristics of the sample. Only two background variables showed significant differences between the development and cross-validation samples: age and length of time with this counselor. An ANOVA analysis revealed

that age is related to the Acceptance subscale for the cross-validation sample ($F(4, 200) = 2.46, P = 0.047$), with younger respondents having significantly higher Acceptance scores. However, for the development sample, age was not significantly related to scores of Acceptance. Length of time with counselor was not significantly related to Acceptance scores for either sample of respondents. Moreover, these two variables were not significantly related to the Rolling with Resistance subscale scores. Thus, the explanation is unclear. If we assume that clients with a longer relationship with their therapists provide a more accurate depiction of the reliability—or more specifically, the problems of reliability—in these scales, we are concerned that they require further development. The reliability of the Acceptance and Rolling with Resistance subscales is thus far inconclusive. We recommend keeping the subscales for the next phase of studies, when we can further investigate whether the reliability of the subscales is dependent on sample characteristics, such as the age of participants.

Inter-rater reliability (consensus scores)

For a client level adherence measure to be considered valid, it is important to demonstrate that there is consensus among different clients' ratings of the same counselor. Inter-rater reliability was computed using the consensus index ($r_{wg(i)}$) developed by James, Demaree and Wolf³¹ for multiple item scales. A higher

**Table 2B.** Confirmatory Factor Analysis of the MET client level fidelity scale.

Item	Loading
Factor 1: Support Self-Efficacy/Elicit Change Talk (SS/ECT)	
My counselor asks for my opinions.	0.75
My counselor helps me to see what I am good at.	0.75
My counselor and I are working toward goals I want to achieve.	0.73
My counselor helps me feel good about positive changes I make	0.72
My counselor and I talk about what I want for my future.	0.72
My counselor believes that I can stay clean and sober	0.71
My counselor respects how I feel.	0.68
My counselor helps me see that I can change if I want to.	0.67
My counselor asks me to talk about reasons why I want to give up drugs or alcohol.	0.62
My counselor helps me see that I am responsible for making changes in my life.	0.55
My counselor and I talk about how my life could be better if I made different choices.	0.52
Factor 2: Avoid Argumentation (AA)	
My counselor argues with me a lot.	0.79
My counselor gets really upset and yells at me if I have a slip.	0.68
My counselor has a hard time understanding how things are done where I come from.	0.61
I feel like I have to defend myself to my counselor.	0.59
When I get upset and mad in sessions, my counselor gets upset and mad, too.	0.54
It seems like my counselor is always angry with me.	0.53
Factor 3: Roll with Resistance (RR)	
My counselor gets mad if I don't follow our treatment plan.	0.56
When I get upset, my counselor tells me that I should stop feeling sorry for myself.	0.56
When we disagree, my counselor tries to talk me into her/his point of view.	0.28
My counselor pushes me to change my life.	0.12
Factor 4: Client Centered Perspective (CC)	
I don't really think that my counselor understands me.	0.68
My counselor mainly talks about his/her own recovery.	0.57
My counselor probably talks about pretty much the same things with all of her/his clients.	0.52
I don't think that my counselor understands what is important to me.	0.52
Factor 5: Express Empathy (EE)	
My counselor tries to see things from my point of view	0.71
My counselor understands why I do things	0.67
I can tell my counselor when I do not think something will work.	0.64
When I have a problem, my counselor listens and helps me come up with my own solutions.	0.56
I can talk to my counselor about what I like about using.	0.52
My counselor understands how hard it is to give up drugs and alcohol.	0.52
My counselor helps me to see how my actions don't always help me meet my goals	0.40
My counselor does not judge me for what I do.	0.38

consensus score implies a high degree of agreement among clients rating the same counselor. James et al³² suggest that $r_{wg(j)}$ scores equal to or larger than 0.70 indicate a good level of consensus among different raters on the same target. As shown in Table 5, for most of the CBT and MET subscales and total scores, the $r_{wg(j)}$ scores are higher than 0.70, except for the Rolling with Resistance (RR) and Client Centered (CC) subscales. Even though the RR and CC subscales have lower consensus scores, they remained in the 0.6 range, which implies a moderate degree of agreement among clients' ratings of the same counselor.

Discussion

This paper reports on the development and initial validation of a client-rated MET-CBT fidelity adherence measure. It is a crucial first step in the development of a cost-effective and consumer-directed fidelity instrument. Results suggest that there is good initial psychometric evidence for a client-rated MET-CBT fidelity measure, with the possible exception of two subscales, and point to the need for further studies of the measure.

For the development sample, Cronbach's alphas were all adequate, $\alpha > 0.70$. However, there were two



Table 3. Descriptive Statistics and Cronbach's alphas for the development and cross-validation samples.

Subscale (# of items)	Development sample			Cross-validation sample		
	M	SD	Alpha	M	SD	Alpha
CBT						
FA/CR (9)	4.13	0.85	0.94	4.13	0.55	0.83
DLS (5)	3.65	0.89	0.79	3.57	0.67	0.67
Total (14)	3.89	0.81	0.93	3.85	0.56	0.85
MET						
SS/ECT (11)	4.13	0.78	0.94	4.21	0.54	0.88
AA (6)	3.82	1.09	0.91	4.06	0.69	0.79
RR (4)	3.07	1.01	0.72	2.98	0.72	0.39
CC (4)	3.59	1.06	0.81	3.67	0.83	0.66
EE (9)	3.91	0.78	0.88	3.95	0.53	0.80
Total (34)	3.70	0.63	0.92	3.95	0.53	0.89

Abbreviations: FA/CR, Functional Analysis/Coping with Risk (CBT1); DLS, Developing Life Skills (CBT2); CBT, Cognitive Behavior Therapy, Total Scale; SS/ECT, Support self-efficacy/Elicit Change Talk (MET1); AA, Avoid Argumentation (MET2); RR, Roll with Resistance (MET3); EE, Express Empathy; (MET4) AC, Acceptance (MET6)? Should we change the label for this new factor?; CC, Client-Centered (MET5); MET, Motivational Enhancement Therapy, Total Scale.

subscales for the cross-validation sample for which the alphas were low: (1) the Roll with Resistance subscale was 0.37, and (2) the Acceptance subscale was 0.57. The large discrepancy of alpha values for the Roll with Resistance subscale and the Acceptance subscale between the development sample (α : RR = 0.72, and AC = 0.71), and cross-validation samples (α : RR = 0.37, and AC = 0.57) could be the result of sampling variation, which would not be a concern. However, it is also possible that clients in the cross-validation sample, with their greater experience with their counselors, present a more accurate and troubling view of these subscales' internal consistency. Therefore, the reliability of the Acceptance and Rolling with Resistance subscales is inconclusive, and should be subject to further testing and development.

We examined the level of consensus, or $r_{wg(j)}$, among clients rating the same counselor on the counselor's MET-CBT skills. For most of the CBT and MET subscales and total scores, the $r_{wg(j)}$ was higher than the acceptable 0.70, except for the Rolling with Resistance and Client Centered subscales. We hypothesize that these two subscales were lower because they tap counselor skills that are more difficult for clients to recognize. Even though the RR and CC subscales had lower consensus scores, they still were in the 0.6 range, which implies a moderate degree of agreement among clients' ratings of the same counselor.

It must be emphasized that the current study was conducted in a real-world context, ie, a community-based addictions treatment agency, where it was

Table 4A. Zero-order correlation among MET-CBT subscales: development sample only.

	FA/CR	DLS	CBT	SS/ECT	AA	RR	EE	CC	MET
DLS	0.75**								
CBT	0.89**	0.89**							
SS/ECT	0.72**	0.77**	0.79**						
AA	0.16**	0.03	0.12*	0.18**					
RR	-0.10*	-0.26**	-0.19**	-0.17**	0.62**				
EE	0.64**	0.73**	0.71**	0.83**	0.11*	-0.20**			
CC	0.17**	-0.01	0.11*	0.11*	0.72**	0.54**	0.03		
MET	0.42**	0.29**	0.39**	0.50**	0.86**	0.63**	0.44**	0.80**	

Notes: **Correlation is significant at the 0.01 level, 2-tailed; *correlation is significant at the 0.05 level, 2-tailed.
Abbreviations: FA/CR, Functional Analysis/Coping with Risk (CBT1); DLS, Developing Life Skills (CBT2); CBT, Cognitive Behavior Therapy, Total Scale; SS/ECT, Support self-efficacy/Elicit Change Talk (MET1); AA, Avoid Argumentation (MET2); RR, Roll with Resistance (MET3); EE, Express Empathy (MET4); CC, Client-Centered (MET5); AC, Acceptance (MET6); MET, Motivational Enhancement Therapy, Total Scale.

**Table 4B.** Zero-order correlation among MET-CBT subscales: cross-validation sample only.

	FA/CR	DLS	CBT	SS/ECT	AA	RR	CC	EA	MET
DLS	0.75**								
CBT	0.90**	0.92**							
SS/ECT	0.75**	0.79**	0.85**						
AA	0.22**	0.21**	0.23**	0.36**					
RR	-0.15*	-0.05	-0.12	-0.01	0.48**				
CC	0.22**	0.23**	0.24**	0.35**	0.67**	0.45**			
EA	0.70**	0.73**	0.79**	0.81**	0.26**	-0.07	0.32**		
MET	0.43**	0.47**	0.49**	0.64**	0.82**	0.59**	0.84**	0.58**	

Notes: **Correlation is significant at the 0.01 level (2-tailed); *correlation is significant at the 0.05 level (2-tailed).

Abbreviations: FA/CR, Functional Analysis/Coping with Risk (CBT-factor 1); DLS, Developing Life Skills (CBT-factor 2); CBT, Cognitive Behavior Therapy, Total Scale; SS/ECT, Support self-efficacy/Elicit Change Talk (MET-factor 1); AA, Avoid Argumentation (MET-factor 2); RR, Roll with Resistance (MET-factor 3); EE, Express Empathy (MET-factor 4); AC, Acceptance (MET-factor 6); CC, Client-Centered (MET-factor 5); MET, Motivational Enhancement Therapy, Total Scale.

not possible to completely control for external variables. Also, methadone clinics are quite different from other addiction treatment agencies, especially since the average length of treatment is considerably longer. Thus, our results may not be generalizable to other community clinics. However, using methadone clinics, which have better attendance and longer-term involvement of clients, helped us to obtain data from clients who were better informants about their clinicians' work. Also, the initial psychometrics for this scale are good; further studies should be designed to provide additional psychometric information on the use of a client-rated MET-CBT adherence measure with different client populations.

Table 5. Consensus ratings for client level MET-CBT measurements.

	Development sample	Cross-validation sample
CBT		
FA/CR	0.76	0.83
DLS	0.66	0.73
Total	0.72	0.80
MET		
SS/ECT	0.78	0.86
AA	0.64	0.78
RR	0.58	0.66
CC	0.64	0.63
EE	0.73	0.82
Total	0.69	0.78

Notes: $r_{wg(i)} = 1 - \frac{\bar{S}_i^2}{S_{mv}^2}$; \bar{S}_i^2 is the obtained average variance of the items in the scale, and S_{mv}^2 is the maximum dissensus distribution ($S_{mv}^2 = 0.5(X_U^2 + X_L^2) - [0.5(X_U + X_L)]^2$); X_U is the upper and X_L is the lower extremes of the response scale.

This study offers some support for the idea of tracking fidelity through consumer surveys, which promises a lower-cost, more objective measurement. Clearly, additional work needs to be done to develop the scale prior to its use for high-stakes purposes such as reimbursement. If successful, we envision coupling validated client-rated adherence measures with newer technology, such as computers with touch screens to regularly receive data from consumers from each agency program. With greater automation, we will be able to develop a system with the ability to insert known valid and reliable EBP items for different centers or programs as training takes place, and to automatically tie both client outcomes (from the management information system) and client-rated adherence ratings with the program and the direct care staff. In this manner, the therapeutic change process will be better understood, ultimately leading to more effective therapeutic techniques and improved consumer outcomes.

Author Contributions

Conceived and designed the experiments: WU, HL, LF, SS. Analysed the data: HL. Wrote the first draft of the manuscript: WU, HL. Contributed to the writing of the manuscript: LF, SS, SHG, KLS, JLK, MO. Agree with the manuscript results and conclusions: All authors. All authors reviewed and approved of the final manuscript.

Funding

This research was supported by a grant, R21DA19781, from the National Institute on Drug Abuse. We are



grateful for the expert supervisor training and consultation provided by Richard Fisher, MSW, and for the participation by agency personnel and clients, who gave of their time and feedback to make this project possible.

Competing Interests

JLK received funding to develop national training resources on MET/CBT 5. Other authors disclose no competing interests.

Disclosures and Ethics

As a requirement of publication author(s) have provided to the publisher signed confirmation of compliance with legal and ethical obligations including but not limited to the following: authorship and contributorship, conflicts of interest, privacy and confidentiality and (where applicable) protection of human and animal research subjects. The authors have read and confirmed their agreement with the ICMJE authorship and conflict of interest criteria. The authors have also confirmed that this article is unique and not under consideration or published in any other publication, and that they have permission from rights holders to reproduce any copyrighted material. Any disclosures are made in this section. The external blind peer reviewers report no conflicts of interest.

References

1. Miller WR, Sorensen JL, Selzer JA, Brigham GS. Disseminating evidence-based practices in substance abuse treatment: A review with suggestions. *Journal of Substance Abuse Treatment*. 2006;31:25–39.
2. Morgenstern J. Effective technology transfer in alcoholism treatment. *Substance Use and Misuse*. 2000;35:1659–78.
3. Beutler LE. The empirically supported treatments movement: A scientist-practitioner's response. *Clinical Psychology: Science and Practice*. 2004; 11(3):225–29.
4. Miller WR, Zweben J, Johnson WR. Evidence-based treatment: Why, what, where, when, and how? *Journal of Substance Abuse Treatment*. 2005;29:267–76.
5. Freemantle N. Implementation strategies. *Family Practice*. 2000;17:S7–10.
6. Gorton TA, Cranford CO, Golden WE, Walls RC, Pawelak JE. Primary care physicians' response to dissemination of practice guidelines. *Archives of Family Medicine*. 1995;4:135–42.
7. Riley KJ, Rieckmann T, McCarty D. Implementation of MET/CBT 5 for adolescents. *Journal of Behavioral Health Services and Research*. 2008; 35(3):304–14.
8. Herschell AD, McNeil CB, McNeil DW. Clinical child psychology's progress in disseminating empirically supported treatments. *Clinical Psychology: Science and Practice*. 2004;11:267–88.
9. Hayes SC, Bissett R, Roget N, Padilla M, Kohlenberg BS, Fisher G. The impact of acceptance and commitment training and multicultural training on the stigmatizing attitudes and professional burnout of substance abuse counselors. *Behavior Therapy*. 2004;35:821–35.
10. Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: A systematic review of 102 trials of interventions to improve professional practice. *Canadian Medical Association Journal*. 1995;153:1423–31.
11. Carroll KM. Treating drug dependence: Recent advances and old truths. In: Miller WR, Heather N, editors. *Treating Addictive Behaviors*. 2nd ed. New York: Plenum Press; 1998.
12. Herschell AD. Fidelity in the field: Developing infrastructure and fine-tuning measurement. *Clinical Psychology: Science and Practice*. 2010;17(3):253–7.
13. Beidas RS, Kendall PC. Training therapists in evidence-based practice: A critical review of studies from a systems-contextual perspective. *Clinical Psychology: Science and Practice*. 2010;17:1–30.
14. Miller WR, Yahne CE, Moyers TB, Martinez J, Pirritano M. A randomized trial of methods to help clinicians learn motivational interviewing. *Journal of Consulting and Clinical Psychology*. 2004;72:1050–62.
15. Carroll KM, Farentinos C, Ball SA, Crits-Christoph P, Libby B, Morgenstern J. MET meets the real world: Design issues and clinical strategies in the Clinical Trials Network. *Journal of Substance Abuse Treatment*. 2002;23:73–80.
16. Schoenwald SK. It's a bird, it's ... fidelity measurement in the real world. *Clinical Psychology: Science and Practice*. 2011;18(2):142–7.
17. McCrady BS, Ziedonis D. American Psychiatric Association practice guidelines for substance use disorders. *Behavior Therapy*. 2001;32:309–36.
18. Finney JW, Moos RH. Psychosocial treatments for alcohol use disorders. In: Nathan PE, Gorman JM, editors. *A Guide to Treatment that Works*. 2nd ed. London: Oxford University Press; 2002.
19. Brown JM, Miller WR. Impact of motivational interviewing on participation and outcome in residential alcoholism treatment. *Psychology of Addictive Behaviors*. 1993;7:211–8.
20. Gentilello LM, Rivera FP, Donovan DM, Jurkovich GJ, Daranciang E, Dunn CW. Alcohol interventions in a trauma center as a means of reducing the risk of injury recurrence. *Annals of Surgery*. 1999;230:473–83.
21. Burke BL, Arkowitz H, Dunn C. The efficacy of motivational interviewing and its adaptations: what we know so far. In: Miller WR, Rollnick S, editors. *Motivational Interviewing: Preparing People for Change*. 2nd ed. New York: Guilford Press; 2002.
22. Miller WR, Rollnick S. *Motivational Interviewing: Preparing People for Change*. 2nd ed. New York: Guilford Press; 2002.
23. Monti PM, Kadden RM, Rohsenow DJ, Cooney NL, Abrams DB. *Treating Alcohol Dependence: A Coping Skills Training Guide*. New York, New York: Guilford Press; 2002.
24. Kahn JH. Factor analysis in counseling psychology research, training, and practice: principles, advances, and applications. *Counseling Psychologist*. 2006;34(5):684–718.
25. Arbuckle JL, Wothke W. AMOS 4.0 user's guide. Chicago: SPSS. 1999.
26. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999;6(1):1–55.
27. Steiger JH. Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*. 1990;25(2):173–80.
28. Hooper D, Coughlan J, Mullen MR. Structural equation modeling: guidelines for determining model fit. *The Electronic Journal of Business Research Methods*. 2008;6:53–60.
29. Akaike H. Factor analysis and AIC. *Psychometrika*. 1987;52(3):317–32.
30. Nunnally JC. *Psychometric Theory*. 2nd ed. New York: McGraw-Hill; 1978.
31. James LR, Demaree RG, Wolf G. Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*. 1984;69:85–98.
32. James LR, Demaree RG, Wolf G. r_{wg} : An assessment of within-group interrater agreement. *Journal of Applied Psychology*. 1993;78(2):306–9.