Festschrift: Gary Bond and Fidelity Assessment

Development and Reliability of a Measure of Clinician Competence in Providing Illness Management and Recovery

Alan B. McGuire, Ph.D.
Laura G. Stull, Ph.D.
Kim T. Mueser, Ph.D.
Meghan Santos, M.S.W.
Abigail Mook, M.S.
Nicole Rose, M.S.
Chloe Tunze, Ph.D.
Laura White, M.S.
Michelle P. Salyers, Ph.D.

Dr. McGuire, Dr. Stull, Ms. Mook, Ms. Rose, Dr. Tunze, Ms. White, and Dr. Salyers are with the ACT Center of Indiana, Department of Psychology, Indiana University–Purdue University Indianapolis, 1481 W. 10th St. (11H), Room D6014, Indianapolis, IN 46202 (abmcguir@iupui.edu). Dr. McGuire is also affiliated with the Roudebush Veterans Affairs Medical Center, Indianapolis, Indiana. Dr. Mueser is with the Center for Psychiatric Rehabilitation, Boston University. Ms. Santos is with the Department of Community and Family Medicine, Geisel School of Medicine at Dartmouth, Dartmouth College, Hanover, New Hampshire. This article is part of a special section, Festschrift: Gary Bond and Fidelity Assessment, for which Dr. Salyers served as guest editor.

Objective: Illness management and recovery (IMR) is an evidence-based, manualized illness self-management program for people with severe mental illness. This study sought to develop a measure of IMR clinician competence and test its reliability and validity. Methods: Two groups of subject matter experts each independently created a clinician-level IMR competence scale based on the IMR Fidelity Scale and on two unpublished instruments used to evaluate provider competence. The two versions were merged, and investigators used the initial version to independently rate recordings of IMR sessions. Ratings were compared and discussed, discrepancies were resolved, and the scale was revised through 14 iterations. The resulting IMR Treatment Integrity Scale (IT-IS) includes 13 required items and three optional items rated only when the particular skill is attempted. Four independent raters then used the IT-IS to score tapes of 60 IMR sessions and 20 control group sessions. Results: The IT-IS showed excellent interrater reliability. A factor analysis supported a one-factor model that showed good internal consistency. The scale successfully differentiated between IMR and control groups. Reliability and validity of individual items varied widely. Conclusions: The IT-IS is a promising

measure of clinician competence in providing IMR. The scale could be used for research and quality assurance and as a supervisory feedback tool. Future research is needed to examine item-level changes, predictive validity of the IT-IS, discriminant validity compared with other more structured interventions, and the reliability and validity of the scale for nongroup IMR. (*Psychiatric Services* 63:XXXXXXXX, 2012; doi: 10.1176/appi.ps.201100144)

Measures of treatment integrity are crucial in determining adherence to a specific program model and differentiating a practice from others (1–4). Treatment integrity is equally important to clinical research and dissemination of evidence-based practices. Accurate dissemination requires specification of the critical components of a given model (5), development of operational definitions of the critical ingredients (6), and development of standardized assessments of the degree of implementation (7). In addition to assisting with dissemination, adherence to program models has been shown to be a predictor of consumer outcomes (7,8). The identification of evidence-based practices, together with the need for broad and accurate dissemination of such practices, has led to a general call for the development of methods to empirically validate adherence to program models (9–13). However, treatment integrity is often neglected in both research and practice (1,14–16).

The illness management and recovery (IMR) program is an evidence-based approach to teaching consumers with severe mental illness strategies for setting and achieving personal recovery goals and acquiring the knowledge and skills to manage their illnesses (17,18). IMR was developed as part of the National Implementing Evidence-Based Practices Project of the Substance Abuse and Mental Health Services Administration (19) and includes an implementation toolkit with a workbook, practice demonstration videos, brochures, and consumer handouts (20). IMR prescribes the use of complex clinical techniques to support consumers in learning the curriculum, setting and achieving goals, and managing their illness. Interventions include cognitive-behavioral techniques (21,22), motivation-based strategies (23), and interactive educational techniques. IMR's effectiveness has been supported by longitudinal (24,25), multisite (26–28), and randomized controlled trials (29–31). Collectively, these studies have shown positive treatment outcomes, with increased illness self-management and coping skills and reduced hospitalization rates (26).

IMR, as with all the programs in the National Implementing Evidence-Based Practices Project (32), utilizes a program-level fidelity scale. The IMR Fidelity Scale (33) is a 13-item scale to assess the degree of implementation of the entire program rather than of an individual practitioner. Each item is rated on a 5-point, behaviorally anchored scale. This scale has been used in state implementation projects (26,28). The sensitivity of the scale has been demonstrated by increased scores after training and consultation (32). This type of program-level scale adopts a broad view of how services are provided by focusing primarily on structural aspects of the program (for example, size of the intervention group) or clinician skills in the aggregate.

Although program-level fidelity is important, integrity is a multipart construct including adherence, competence, and differentiation (16). Therapist competence—"level of skill shown by the therapist in delivering the treatment" (34)—may be more appropriate for programs requiring specific, nuanced clinical interventions. For instance, practitioners must understand the program model and have the skills to implement it (8,35,36). Moreover, treatment integrity measures that assess actual use of knowledge and skills are crucial and are an advancement over assessment techniques that rely on self-report or chart review. Various evidence-based psychotherapy models, such as cognitive-behavioral therapy for psychosis (37,38), multisystemic therapy (39), and adolescent substance abuse treatment (40), often assess treatment integrity at the clinical interaction level. Integrity at this level has been linked to positive outcomes for consumers (41). With the exception of some measures focused on practitioners' general knowledge (42–45), evidence-based psychiatric rehabilitation programs lack clinician-level instruments, particularly those addressing competence. Therefore, the goal of the project described in this report was to create a clinician-level competency scale for IMR and test its psychometric properties.

Methods

Creation of the scale

Two groups of subject matter experts (group 1: ABM, LGS, and MPS; group 2: KTM and MS) each independently created clinician-level IMR competence scales based on the IMR Fidelity Scale (33) and on two unpublished instruments used to evaluate provider competence: the Minnesota IMR Clinical Competence Scale and the IMR Knowledge Test. The scale creators included an IMR model creator, three researchers with extensive experience in IMR implementation, and an experienced IMR clinical supervisor. The scale was developed to rate either live or audio-recorded sessions. The draft scales from the two groups included 15 and 24 items, respectively. The content of these items overlapped; however, one group included an item on recovery orientation (retained in the final version) and items pertinent to specific IMR modules (eliminated from the final version). Items were compared and reconciled.

We initially set out to create behaviorally specific anchors using a 5-point scale, to be consistent with the IMR Fidelity Scale. For five items we were able to create anchors with good agreement between investigators. However, we had difficulty specifying anchors for 11 of the items, and we used a more generic set of anchors for those items that was based on the work of Blackburn and colleagues (46). Higher scores correspond to greater competence in administering the target IMR element. After the initial version was created, investigators independently rated recordings of IMR sessions and then compared and discussed ratings, rationale for ratings, and any discrepancies. Investigators revised the scale, anchors, and rating criteria on the basis of these discussions. This process was repeated until no major rating discrepancies existed and investigators agreed that no further changes were necessary. This process took 14 iterations. The resulting IMR Treatment Integrity Scale (IT-IS) includes 13 required items and three optional items rated only when the particular skill is attempted. [The scale is available online as a data supplement to this article.]

Reliability testing

Sampling. Group sessions were audiotaped as part of a randomized trial of IMR for adults with schizophrenia. Participants were randomly assigned to either weekly IMR or control groups. Participants included veterans at the local Department of Veterans Affairs medical center, as well as consumers at a local community mental health center, with groups held at the respective facilities. Inclusion criteria were a diagnosis of schizophrenia or schizoaffective disorder confirmed by the Structured Clinical Interview for DSM-IV (SCID) (47), passing scores on a simple cognitive screen (48), willingness to participate in a group intervention, and current enrollment in services at either participating facility. Participants were paid $20 for each completed interview, but they were not paid for group attendance.

Group facilitators. IMR and control groups were facilitated initially by a licensed clinical social worker and later by a doctoral-level clinical psychologist (ABM), both of whom had previous experience providing and consulting on IMR. Groups were cofacilitated by clinical psychology graduate students (including LGS, AM, NR, CT, and LW) with a range of clinical experience.

Procedures. IMR groups were conducted according to the group guidelines in the revised IMR implementation toolkit (49). A typical group included brief socialization, discussion of items for the agenda, review of previously covered material, presentation of new material (guided by educational handouts from the IMR workbook and taught using motivation-based, educational, and cognitive-behavioral techniques), and goal setting and follow-up. The control group consisted of unstructured support, in which group members chose discussion topics and facilitators encouraged participation and maintained basic group ground rules (for example, mutual respect and confidentiality). Facilitators were instructed not to use educational materials or IMR-related teaching techniques, such as cognitive-behavioral techniques, but rather to use supportive therapeutic techniques (for example, active listening).

For rating we randomly selected 60 IMR sessions and 20 control group sessions (without replacement, separate draws for IMR and control). Raters were four clinical psychology graduate students (AM, NR, CT, and LW) with experience providing IMR. Raters were trained by the scale creators; scale creators did not serve as raters. Each rater scored 30 IMR tapes and ten control group tapes, and each session was rated by two raters (each rater overlapped with every other rater on an equal number of sessions). Both raters independently scored each session and then discussed any discrepancies and reached a consensus rating.

This study was conducted between May 2009 and May 2011 and was approved by the Indiana University–Purdue University Indianapolis Institutional Review Board.

Analyses. First we examined interrater reliability for the total IT-IS and individual items by computing intraclass correlation coefficients (ICCs); sessions in which a rater was also a group cofacilitator (18 sessions, or 23%) were excluded from this analysis. Systematic rater bias was assessed by comparing mean ratings of total scores across raters by using

4

analyses of variance. The remainder of the analyses were conducted with one of the two raters' scores, randomly selected (except when one rater was a group cofacilitator, in which case the other rater's scores were used).

The second set of analyses examined the factor structure and internal consistency of the scale. We considered two potential factor structures—a one-factor structure and a two-factor structure (general therapeutic elements and IMR-specific elements). We used confirmatory factor analyses to compare the relative fit of the two models (50,51) and selected the best model on the basis of goodness of fit and parsimony. Descriptive statistics were examined for range, distribution, and ceiling-floor effects for individual items. Internal consistency was examined through inspection of item-to-total correlations and Cronbach's alpha. Given the meaningful difference between adherence (using prescribed elements) and competence (skill in using the elements), item analyses were repeated with just the IMR sessions (excluding the control sessions) to test the relationship between skill at specific IMR components within sessions.

We also evaluated whether IMR competence was independent of when and where the session took place. To examine this, we correlated group date and IT-IS score and conducted an independent-samples t test comparing IT-IS scores by site. Finally, construct validity was examined through comparison of known groups. We hypothesized that IMR sessions would receive higher scores than control group sessions on the total scale and on the specific IMR subscales. We tested this hypothesis by using independent-samples t tests and excluding sessions in which a rater was also the group leader (given leaders' a priori knowledge of the group condition).

Results

As shown in Table 1, interrater reliability for the total scale score was excellent (ICC=.92, p<.001). Reliability for individual items was variable, with ICCs ranging from .03 (group member involvement) to .95 (IMR curriculum). Interrater reliability tended to be worse for general therapeutic items and items using the generic anchors. Interrater reliability did not vary substantially by rater pair (Table 2). Mean scores did not differ by rater: rater 1, 3.33±.82; rater 2, 3.38 ±.92; rater 3, 3.27 ±.77; rater 4, 3.41±.75). The remaining analyses were conducted using one of the two raters' scores.

Total and item means, standard deviations, and ranges are presented in Table 3. In regard to the confirmatory factor analysis, goodness-of-fit indicators were compared between the one-factor and two-factor models. Values were as follows for the one-factor model: $\chi2$=167.78, df=64; standardized root mean square residual (RMSR)=.08; adjusted goodness-of-fit index (GFI)=.71; root mean square error of approximation (RMSEA)=.13; and Bentler comparative fit index=.88. For the two-factor model, these values were as follows: $\chi2$=244.03, df=63; standardized RMSR=.26; adjusted GFI=.67; RMSEA=.17; and Bentler comparative fit index=.78. Overall, results indicate that the one-factor model fit is a good fit and better than the two-factor model.

Regarding internal consistency, the items specific to group processes (group member involvement and enlisting mutual support) were not calculated for 18 sessions (23%)

5

because those sessions were attended by only one participant. Therefore, alphas and item-to-total correlations were calculated for the 62 sessions for which all items could be calculated. Overall internal consistency was high both including and excluding the group items (Cronbach's {α alpha}=.90 and .91, respectively). As shown in Table 1, item-to-total correlations ranged from very poor (enlisting mutual support, r=−.07) to very strong (structure and efficient use of time, r=.91). Alphas did not increase more than .02 with the removal of any item. When only the IMR session tapes were used, the overall internal consistency was weaker but still in the acceptable range ({α alpha}=.79). Individual items varied in item-to-total correlations; enlisting mutual support was negatively related to the total score.

As hypothesized, total IT-IS scores were significantly higher for IMR sessions than for control group sessions (Table 3), supporting the construct validity of the scale. Means for items were higher for IMR sessions than for control sessions, with the exception of enlisting mutual support. The scores for the VA sessions (N=44, mean score=3.35±.86) did not differ from those for the sessions conducted at the community site (N=36, mean score=3.36±.81). The mean IT-IS score was also not related to date of the group session.

Discussion

This study provided promising preliminary results on the reliability and validity of the IT-IS, a newly developed measure of IMR clinician competence. The total scale had excellent interrater reliability and adequate item-level interrater reliability, despite some notable exceptions. The scale was internally consistent and best fit a one-factor model. As hypothesized, scores on the IT-IS did not systematically differ by treatment site, rater, or date of the session.

Two items are candidates for removal from the IT-IS. The item about enlisting mutual support did not distinguish IMR sessions from control group sessions and was negatively related to overall competence within IMR sessions. Conversely, therapeutic relationship did differ between the IMR and control groups, although it theoretically should not. Interrater reliability was also poor for this item, and it was only weakly related to the total score in the total sample. There was also a ceiling effect (mean=4.5) in the sample, which could account for the low reliability and minimal relationship to total score. Therapeutic alliance is an important common factor in psychotherapy and has been found to moderate the link between treatment integrity and treatment outcomes in other research (52,53). It is possible that wider variability of this item would be found if a larger number of group leaders were sampled. The scoring criteria for this item could also be revised to make it less lenient. Another alternative would be to supplement the IT-IS with other validated measures of common therapeutic factors (for example, the Working Alliance Inventory [54]).

The construct validity of the scale was supported: IMR sessions were rated higher than control group sessions on the total IT-IS and on all but one IT-IS item. Structure and efficient use of time was most strongly related to total score and may be most indicative of an IMR session compared with an unstructured intervention group. However, this relationship may be less predictive of IMR when compared with other structured

6

interventions, such as cognitive-behavioral therapy or cognitive-processing therapy. Future work is needed to examine how the scale performs when IMR is compared with other structured interventions. Within the IMR sample, action planning was most strongly associated with the total score and therefore may be most indicative of quality within IMR. In our clinical experience with IMR, assisting consumers to convert abstract recovery goals into tangible steps may be emblematic of IMR.

Implications
The IT-IS has several potential important uses. Clinically, the IT-IS can serve as a valuable training tool by specifying the critical elements of IMR and the behaviors associated with competent practice. In addition, the IT-IS can be used by supervisors as a tool to provide specific feedback and develop performance improvement goals (55). From a programmatic standpoint, the IT-IS provides an indicator of variability between clinicians and could be used to assess and plan for additional training needs (56).

In research applications, previous studies have used the IMR Fidelity Scale to monitor overall program fidelity. However, there may be significant variability in the quality of IMR implementation between practitioners within the same program. By monitoring competence at the level of clinical interactions, researchers can obtain a more fine-grained evaluation of the IMR process. An important next step will be examining the relationship between competence (IT-IS items), IMR mechanism of action, and outcomes.

Monitoring with both the IMR Fidelity Scale and the IT-IS would ensure faithful implementation and provide accountability for funding, regulatory bodies, and other stakeholders. The IT-IS complements and extends the IMR Fidelity Scale by assessing individual clinicians. However, there are some structural elements on the IMR Fidelity Scale that are not captured by the IT-IS (for example, group size, session length, and comprehensive curriculum) that would still be important to assess at the program level. We view the two scales as complementary, rather than competing, tools.

Limitations
This study had several limitations. The study used a fairly homogeneous sample of sessions, and the group sessions were cofacilitated by creators of the IT-IS or IT-IS raters (although only sessions rated by two raters who were not cofacilitators in the session were included in interrater reliability analyses). Although session recordings were blinded to the treatment condition and date of the recording, the raters' knowledge of the study and other treatment providers made true blinding impossible. Future studies should use raters who play no role in the intervention. In a related issue, the highly structured and technically specific nature of IMR makes it highly distinguishable from an unstructured control group. Future studies should use the IT-IS to compare IMR to other structured, but distinct, interventions, such as cognitive-behavioral training or social skills training, to provide a more rigorous test of known-groups validity. The IT-IS should also be evaluated in individual IMR sessions to ensure applicability across modalities. Finally, the IT-IS included three optional items that are rated only when the specific element is used. In the study sample, application of these elements was not

sufficient to permit analysis of these items. Future studies should intentionally sample for sessions in which these elements could be rated.

Conclusions

The IT-IS is a promising measure of the competence of IMR practitioners. A revised version of the scale could be valuable in a variety of contexts including clinical, supervisory, administrative, and research. Future studies are needed to replicate this study with independent raters and to extend the research to nongroup IMR.

References

1.   Moncher FJ, Prinz RJ: Treatment fidelity in outcome studies. Clinical Psychology Review 11:247–266, 1991
2.   Bond GR: Introduction to a special section: measurement of fidelity in psychiatric rehabilitation research. Mental Health Services Research 2:73, 2000
3.   Mowbray CT, Holder M, Teague GB, et al: Fidelity criteria: development, measurement, and validation. American Journal of Evaluation 24:315–340, 2003
4.   Teague GB, Bond GR, Drake RE: Program fidelity in assertive community treatment: development and use of a measure. American Journal of Orthopsychiatry 68:216–232, 1998
5.   Bickman L: The functions of program theory; in Using Program Theory in Evaluation. Edited by Bickman L. San Francisco, Jossey-Bass, 1987
6.   Bond GR: Variations in an assertive outreach model. New Directions for Mental Health Services 52:65–80, 1991
7.   McGrew JH, Bond GR, Dietzen LL, et al: Measuring the fidelity of implementation of a mental health program model. Journal of Consulting and Clinical Psychology 62:670–678, 1994
8.   Drake RE, Bond GR, Rapp C: Explaining the variance within supported employment programs: comment on "What predicts supported employment outcomes?" Community Mental Health Journal 42:315–318, 2006
9.   Brekke JS: The model-guided method for monitoring program implementation. Evaluation Review 11:281–299, 1987
10.   Drake RE, Osher FC, Wallach MA: Homelessness and dual diagnosis. American Psychologist 46:1149–1158, 1991

11.   Levine M, Toro P, Perkins D: Social and community interventions. Annual Review of Psychology 44:525–558, 1993

12.   Mowbray CT: Community treatment for the seriously mentally ill: is this community psychology? American Journal of Community Psychology 18:893–902, 1990

13.   Olfson M: Assertive community treatment: an evaluation of the experimental evidence. Hospital and Community Psychiatry 41:634–641, 1990

14.   Brekke JS: What do we really know about community support programs? Strategies for better monitoring. Hospital and Community Psychiatry 39:946–952, 1988

15.   Borrelli B, Sepinwall D, Ernst D, et al: A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. Journal of Consulting and Clinical Psychology 73:852–860, 2005

16.   Perepletchikova F, Treat TA, Kazdin AE: Treatment integrity in psychotherapy research: analysis of the studies and examination of the associated factors. Journal of Consulting and Clinical Psychology 75:829–841, 2007

17.   Mueser KT, Corrigan PW, Hilton DW, et al: Illness management and recovery: a review of the research. Psychiatric Services 53:1272–1284, 2002

18.   Mueser KT, Meyer PS, Penn DL, et al: The illness management and recovery program: rationale, development, and preliminary findings. Schizophrenia Bulletin 32:32–43, 2006

19.   Drake RE, Goldman HH, Leff HS, et al: Implementing evidence-based practices in routine mental health service settings. Psychiatric Services 52:179–182, 2001

20.   Mueser KT, Gingerich S (eds): Illness Management and Recovery Implementation Resource Kit. Rockville, Md, Center for Mental Health Services, Substance Abuse and Mental Health Services Administration, 2002

21.   Kelly GR, Scott JE: Medication compliance and health education among outpatients with chronic mental disorders. Medical Care 28:1181–1197, 1990

22.   Gould RA, Mueser KT, Bolton E, et al: Cognitive therapy for psychosis in schizophrenia: an effect size analysis. Focus 2:95–101, 2004

23.   Arkowitz H, Westra HA, Miller WR, et al (eds): Motivational Interviewing in the Treatment of Psychological Problems. New York, Guilford, 2008

24.   Roe D, Hasson-Ohayon I, Salyers MP, et al: A one year follow-up of illness management and recovery: participants' accounts of its impact and uniqueness. Psychiatric Rehabilitation Journal 32:285–291, 2009

25.   Salyers MP, Hicks LJ, McGuire AB, et al: A pilot to enhance the recovery orientation of assertive community treatment through peer provided illness management and recovery. American Journal of Psychiatric Rehabilitation 12:191–204, 2009

26.   Salyers MP, McGuire AB, Rollins AL, et al: Integrating assertive community treatment and illness management and recovery for consumers with severe mental illness. Community Mental Health Journal 46:319–329, 2010

27.   Salyers MP, Rollins AL, Clendenning D, et al: Impact of illness management and recovery programs on hospital and emergency room use by Medicaid enrollees. Psychiatric Services 63:509–515, 2012

28.   Salyers MP, Godfrey JL, McGuire AB, et al: Implementing the illness management and recovery program for consumers with severe mental illness. Psychiatric Services 60:483–490, 2009

29.   Hasson-Ohayon I, Roe D, Kravetz S: A randomized controlled trial of the effectiveness of the illness management and recovery program. Psychiatric Services 58:1461–1466, 2007

30.   Levitt AJ, Mueser KT, DeGenova J, et al: Randomized controlled trial of illness management and recovery in multiple-unit supportive housing. Psychiatric Services 60:1629–1636, 2009

31.   Färdig R, Lewander T, Melin L, et al: A randomized controlled trial of the illness management and recovery program for persons with schizophrenia. Psychiatric Services 62:606–612, 2011

32.   McHugo GJ, Drake RE, Whitley R, et al: Fidelity outcomes in the National Implementing Evidence-Based Practices Project. Psychiatric Services 58:1279–1284, 2007

33.   Mueser KT, Gingerich S, Bond GR, et al: Illness management and recovery fidelity scale; in Illness Management and Recovery Implementation Resource Kit. Edited by Mueser KT, Gingerich S. Rockville, Md, Center for Mental Health Services, Substance Abuse and Mental Health Services Administration, 2002

34.   Waltz J, Addis ME, Koerner K, et al: Testing the integrity of a psychotherapy protocol: assessment of adherence and competence. Journal of Consulting and Clinical Psychology 61:620–630, 1993

35.   Baer JS, Ball SA, Campbell BK, et al: Training and fidelity monitoring of behavioral interventions in multi-site addictions research. Drug and Alcohol Dependence 87:107–118, 2007

36.   Fixsen DL, Naoom SF, Blase KA, et al: Implementation Research: A Synthesis of the Literature. FMHI pub 231. Tampa, Fla, University of South Florida, Louis de la Parte Florida Mental Health Institute, the National Implementation Research Network, 2005

37.   Startup M, Jackson M, Pearce E: Assessing therapist adherence to cognitive-behaviour therapy for psychosis. Behavioural and Cognitive Psychotherapy 30:329–339, 2002

38.   Haddock G, Devane S, Bradshaw T, et al: An investigation into the psychometric properties of the Cognitive Therapy Scale for Psychosis (CTS-PSY). Behavioural and Cognitive Psychotherapy 29:221–233, 2001

39.   Schoenwald SK, Henggeler SW, Brondino MJ, et al: Multisystemic therapy: monitoring treatment fidelity. Family Process 39:83–103, 2000

40.   Hogue A, Dauber S, Chinchilla P, et al: Assessing fidelity in individual and family therapy for adolescent substance abuse. Journal of Substance Abuse Treatment 35:137–147, 2008

41.   Henggeler SW, Melton GB, Brondino MJ, et al: Multisystemic therapy with violent and chronic juvenile offenders and their families: the role of treatment fidelity in successful dissemination. Journal of Consulting and Clinical Psychology 65:821–833, 1997

42.   Casper ES, Oursler J, Schmidt LT, et al: Measuring practitioners' beliefs, goals, and practices in psychiatric rehabilitation. Psychiatric Rehabilitation Journal 25:223–234, 2002

43.   Casper ES, Oursler JD: The psychiatric rehabilitation beliefs, goals, and practices scale: sensitivity to change. Psychiatric Rehabilitation Journal 26:311–314, 2003

44.   Chinman M, Young AS, Rowe M, et al: An instrument to assess competencies of providers treating severe mental illness. Mental Health Services Research 5:97–108, 2003

45.   Salyers MP, Rollins AL, Bond GR, et al: Development of a scale to assess practitioner knowledge in providing integrated dual disorders treatment. Administration and Policy in Mental Health and Mental Health Services Research 34:570–581, 2007

46.   Blackburn I, James IA, Milne DL, et al: The Revised Cognitive Therapy Scale (CTS-R): psychometric properties. Behavioural and Cognitive Psychotherapy 29:431–446, 2001

47.   First M, Spitzer R, Gibbons M, et al: Structured Clinical Interview for DSM-IV Axis I Disorders, Clinician Edition (SCID-CV). New York, New York State Psychiatric Institute, Biometrics Research, 1997

48.   Callahan CM, Unverzagt FW, Hui SL, et al: Six-item screener to identify cognitive impairment among potential subjects for clinical research. Medical Care 40:771–781, 2002

49.   Gingerich S, Mueser KT (eds): Illness Management and Recovery Implementation Resource Kit: Group Facilitators Guide. Rockville, Md, Center for Mental Health Services, Substance Abuse and Mental Health Services Administration; 2002

50.   Sun J: Assessing goodness of fit in confirmatory factor analysis. Measurement and Evaluation in Counseling and Development 37:240–257, 2005

51.   Cokley KO, Bernard N, Cunningham D, et al: A psychometric investigation of the Academic Motivation Scale using a United States sample. Measurement and Evaluation in Counseling and Development 34:109–119, 2001

52.   Barber JP, Triffleman E, Marmar C: Considerations in treatment integrity: implications and recommendations for PTSD research. Journal of Traumatic Stress 20:793–805, 2007

53.   Feeley M, DeRubeis RJ, Gelfand LA: The temporal relation of adherence and alliance to symptom change in cognitive therapy for depression. Journal of Consulting and Clinical Psychology 67:578–582, 1999

54.   Horvath A, Greenberg LS: Development and validation of the Working Alliance Inventory. Journal of Counseling Psychology 36:223–233, 1989

55.   Breitenstein SM, Gross D, Garvey CA, et al: Implementation fidelity in community-based interventions. Research in Nursing and Health 33:161–173, 2010

56.   Martino S, Canning-Ball M, Carroll KM, et al: A criterion-based stepwise approach for training counselors in motivational interviewing. Journal of Substance Abuse Treatment 40:357–365, 2010