The TMACT: A New Tool for Measuring Fidelity to Assertive Community Treatment

Maria Monroe-DeVita, Gregory B. Teague and Lorna L. Moser

Journal of the American Psychiatric Nurses Association 2011 17: 17
DOI: 10.1177/1078390310394658

Journal of the American Psychiatric Nurses Association
17(1) 17–29

DOI: 10.1177/1078390310394658
http://japna.sagepub.com
SAGE

The TMACT: A New Tool for Measuring Fidelity to Assertive Community Treatment

Maria Monroe-DeVita[1], Gregory B. Teague[2], and Lorna L. Moser[3]

Abstract

BACKGROUND: Fidelity assessment is important for implementation of evidence-based practices (EBPs), including assertive community treatment (ACT). **OBJECTIVES:** The TMACT, an enhanced fidelity tool, was developed and pilot-tested to better assess critical ACT structures and processes. **DESIGN:** Ten ACT teams were administered the TMACT and the long-standing ACT fidelity measure, the Dartmouth Assertive Community Treatment Scale (DACTS), at baseline, 6, 12, and 18 months. **RESULTS:** Overall, fidelity scores for all 10 teams were relatively high. Six teams showed improvement, concluding with high TMACT scores at 18 months. Four teams with significantly lower total scores had experienced turnover and organizational barriers. TMACT ratings were higher in core ACT practices than in recovery practices and EBPs. TMACT scores rose steadily but were significantly lower than DACTS scores, which remained unchanged. **CONCLUSIONS:** The TMACT sets higher performance standards through enhanced assessment of recovery-orientation, EBPs, and teamwork and is more sensitive to change than the DACTS.

Keywords

assertive community treatment, program for assertive community treatment, ACT/PACT, schizophrenia, quality improvement, evidence-based practices

Since its original implementation in Madison, Wisconsin, approximately 30 years ago (Stein & Test, 1980), the assertive community treatment (ACT) model has evolved in many ways to fit with the changing philosophy and practice of the recovery movement as well as the integration of evidence-based and other state-of-the-art practices (Bond, Salyers, Rollins, Rapp, & Zipple, 2004; Salyers & Tsemberis, 2007). Increased attention toward integrating these important areas within ACT speaks to the need for a fidelity tool that captures these core processes. The pre-existing ACT fidelity tool—the Dartmouth Assertive Community Treatment Scale (DACTS; Teague, Bond, & Drake, 1998)—has not only been outdated in relation to measuring these evolving practices but also omits other important core ACT team *processes* (vs. simply articulating team *structure*) hypothesized to be essential to high functioning ACT programs. This article provides an overview of the new and improved ACT fidelity measure—the Tool for Measurement of Assertive Community Treatment (TMACT). In particular, the authors describe the following: (a) what fidelity is and why it is important; (b) a history of ACT fidelity measurement, including a description and summary of limitations of the DACTS; (c) an overview of the development of the TMACT; (d) a description of the TMACT scale and fidelity assessment process; (e) results from piloting the TMACT in one state; and (f) a discussion of next steps regarding ongoing TMACT research and development.

What Is Fidelity and Why Is It Important?

Program fidelity refers to the extent to which a program adheres to the intended model, both including features that are critical to achieving the intended outcomes and excluding those that would interfere (Waltz, Addis, Koerner, & Jacobson, 1993). Measures of fidelity typically

[1]Maria Monroe-DeVita, PhD, University of Washington, Seattle, WA, USA
[2]Gregory B. Teague, PhD, University of South Florida, Tampa, FL, USA
[3]Lorna L. Moser, PhD, Duke University, Durham, NC, USA

Corresponding Author:
Maria Monroe-DeVita, University of Washington School of Medicine, 2815 Eastlake Avenue E., Suite 200, Seattle, WA 98102, USA
Email: mmdv@u.washington.edu

include a combination of qualitative and quantitative indicators that assess the key features of an intervention as implemented relative to the optimal or evidence-based model (Mowbray, Holder, Teague, & Bybee, 2003). To assess fidelity using a standardized instrument, a program model or service must be clearly defined, explicating, for example, the population to be served, staffing and roles, and specific program processes and interventions (Teague & Monroe-DeVita, in press).

The value of program fidelity measurement spans across clinical, research, and administrative purposes (Bond, Evans, Salyers, Williams, & Kim, 2000; Teague, Drake, & Ackerson, 1995; Teague et al., 1998) and rests on the relationship between high fidelity and desired outcomes. Indeed, studies generally demonstrate that high fidelity is tied to better outcomes across many evidence-based practices (EBPs), including ACT (Becker, Smith, Tanzman, Drake, & Tremblay, 2001; McHugo, Drake, Teague, & Xie, 1999).

Many outcomes, especially those that reflect improvement in quality of life or rehabilitative areas, typically take a long time to show results, thereby delaying meaningful feedback. In some ways, fidelity measurement can serve as a proxy for outcome measurement. Thus, assessing whether a program is adhering to the model can provide interim information about whether it will demonstrate its desired outcomes (Teague & Monroe-DeVita, in press).

In addition to serving as a proxy for outcome measurement, ACT fidelity has been linked to systemwide cost savings. Latimer (1999) found that higher fidelity ACT programs serving consumers with a recent history of high service utilization were more cost-effective, which is a typical administrative concern.

Service providers can use fidelity measures as a clinical yardstick, comparing actual with intended practice. New program providers may use the fidelity measure as a guideline to ensure replication, whereas existing programs may refer to such measures to help prevent program drift. In research, fidelity assessment is critical; without knowing exactly what practice was implemented, it is very difficult to attribute changes or differences to the practice (McHugo et al., 1998). Fidelity measurement allows one to deconstruct the model to better identify which ingredients are most critical in terms of outcomes of interest. Clearly defined model fidelity provides a conceptual base for informed adaptation and innovation, all of which can be empirically studied. Furthermore, fidelity assessment provides a measurement of strength in multisite studies and may provide a basis for program inclusion and exclusion in such studies.

Perhaps most important for the purposes of our work, fidelity assessment has the potential for providing valuable information to guide ongoing performance improvement; because fidelity measures delineate the specific structure and processes that are either shown or believed to lead to better outcomes for persons served, they are in essence a measure of *program quality*.

Previous Measurement of ACT Fidelity

Initially, evaluation of ACT program quality consisted of training, consultation, and follow-up site visits by acknowledged experts. As the number of ACT teams in the United States grew, the field increasingly recognized that more rigorous, standardized evaluation approaches to ensure adequate replication and model adherence were needed (Moncher & Prinz, 1991). Brekke and his colleagues (Brekke, 1988; Brekke & Test, 1992) showed that empirical methods could be used to examine program dimensions and differentiate program models. Shortly thereafter, McGrew, Bond, Dietzen, and Salyers (1994) surveyed ACT experts to identify critical ingredients of the intervention, then incorporated the results into a 17-item scale, and retrospectively rated 18 programs. They found that an important outcome, reduction in hospital use, was significantly correlated with shared caseloads, total number of contacts, 24-hour availability, a nurse on the team, and daily team meetings.

Informed by the work of McGrew et al. (1994) as well as by personal communication with developers and leading practitioners of the model, Drake and colleagues (Drake et al., 1998; Teague et al., 1995) developed a measure to use in a multisite trial of an adaptation of ACT for persons with co-occurring psychiatric and addictive disorders. Data were collected from multiple sources, such as on-site observations, interviews, and review of administrative and clinical records, to inform ratings on a set of anchored, 5-point scales. Each item represented a feature of the program deemed important to faithful implementation of the intended model. Decisions about content were based on expert consensus about critical features of the model as well as practicality and need within the context of the parent study. A positive relationship was found between fidelity and both retention in treatment and reduction in hospital admissions (McHugo et al., 1999).

This initial measure (Teague et al., 1995) was subsequently modified, in part with the addition of several items based on emerging expert consensus, including some that had been validated in an earlier fidelity study (McGrew et al., 1994). The new measure, the DACTS (Teague et al., 1998), had 26 items, each rated on its own 5-point scale representing the range from not implemented at all—equivalent to standard care—to fully implemented, and a separate qualitatively (e.g., description of the team's role in offering crisis services) or quantitatively (e.g., average number of minutes spent with a consumer in a given week) descriptive anchor specified for each point.

Despite acknowledgment by the authors that it omitted some important features, the DACTS soon became the standard fidelity measure for ACT (Phillips et al., 2001), in part because it became available in advance of the ACT program manual (Allness & Knoedler, 1998) and in part because of its helpful protocol and visually clear layout, which made it a convenient guide for program implementation. The development approach and structure of the measure were incorporated into general practice for developing fidelity measures in psychiatric rehabilitation research (Bond et al., 2000) and served as the model for corresponding measures for other types of interventions in the National Implementing EBPs Project (McHugo et al., 2007) and their respective Substance Abuse and Mental Health Services Administration (SAMHSA) toolkits (SAMHSA, 2010). Salyers et al. (2003) identified preliminary criteria for adequate performance using the DACTS, and some states began using it to guide funding and certification.

Nonetheless, there were shortcomings. Several hypothetically critical ingredients were either assessed obliquely or knowingly omitted altogether (e.g., team functioning, assessment, and treatment planning). Treatment practices in the field came to embrace both the broader framework of recovery and provision of EBPs; the ACT model evolved along with the field, and over time the gap between the DACTS and optimal ACT increased (Salyers & Tsemberis, 2007). Rating anchors for some items needed adjustment, and more guidance in administration of the tool was needed both to improve validity and reliability and to support use in consultation and performance improvement.

The DACTS also shared in a more general problem among other fidelity measures: emphasis on structure over process. In the Donabedian (1988) model of program quality, *structures* (e.g., staffing, organization) of a program provide a platform for the *processes* (e.g., interactions, procedures) operating within a program or intervention, which in turn yield the desired *outcomes*. Mowbray et al. (2003) argued that an emphasis on structural features would limit a fidelity measure's effectiveness; processes are closer to outcomes, and critical ingredients are most especially critical processes.

These shortcomings in the DACTS could be expected to influence research, practice, and administration/policy. Fidelity measures can inevitably tap only selected aspects of an intervention, serving at best as a coarse model of critical ingredients. The DACTS began as a research tool in a specific research setting; program features could be measured as needed without concern for possible effects on the programs. In the real world, feedback from fidelity measures can affect programs by focusing greater attention on features that are assessed. To the extent that clinical processes critical to outcomes are omitted from a fidelity measure, they may be less frequently or less faithfully carried out, and incentives such as accreditation or funding linked to fidelity scores could further exacerbate the imbalance of program elements. The resulting distortion to a given service model could undermine its effectiveness, resulting in lost opportunity for consumers, wasted resources for the system, and less valid research conclusions about effectiveness.

The DACTS has arguably made a valuable contribution over the years, better focusing efforts on high-fidelity ACT implementation as well as fidelity instrument design and general use. However, it has changed little in the past decade and a half and has not evolved along with changes in the environments in which it has been used. Anecdotal and other evidence reinforces concern about the limitations described (Dixon et al., 2010; Salyers & Tsemberis, 2007). An upgrade to ACT fidelity measurement has been due.

Development of the TMACT

The implementation of 10 new ACT teams in the State of Washington in 2007 provided the impetus for development of an enhanced version of the DACTS. Washington State richly funded these 10 new teams, providing $2.2 million for development and training in state fiscal year 2007, with an additional $10.4 million per year. This amounted to approximately $1.4 million per full team (serving 90-100 consumers) and $650,000 per half team (serving 42-50 consumers; Bjorklund, Monroe-DeVita, Reed, Toulon, & Morse, 2009). At the time, the Washington State mental health authority had decided to adopt and adapt the National Program Standards for ACT teams (Allness, 2003). Many of those adaptations included the need to directly integrate person-centered and recovery-oriented practices to address consumers' and advocates' concerns about potentially implementing a coercive and/or paternalistic version of ACT. To further address these concerns, teams received extensive training and consultation in recovery-oriented services, and many other EBPs. The need for a more comprehensive tool to assess fidelity to these increasingly essential elements of ACT quickly became apparent, as well as the implications for such a tool beyond Washington State ACT teams.

To proceed with this major effort, the authors opted to not reinvent the wheel but to preserve the current DACTS template, knowing that it captured many elements that are indeed critical to ACT, has been pilot-tested across 10 years of program evaluation and research efforts, and provided a sound protocol template. Furthermore, the 5-point behaviorally anchored scale was a useful framework, particularly for application of the tool for performance improvement purposes.

The TMACT has been developed over the course of 3 1/2 years of continuous consultation with experts, including academic researchers, providers, administrators, and ACT consultants and evaluators, as well as feedback from multistate pilot administration efforts. Initially, the authors worked from DACTS revisions initiated by colleagues at the ACT Center of Indiana, who shared potential revisions based on their extensive experience with the tool, and engaged several national experts on ACT and related areas of expertise (e.g., person-centered planning, integrated dual disorder treatment). The authors then incorporated feedback from Washington State ACT stakeholders, including collaborators from the Division of Behavioral Health and Recovery, and piloted a 52-item version of the enhanced fidelity scale within Washington State. Further iterations were guided by the authors' experience in either directly conducting or overseeing a total of 75 TMACT fidelity reviews as of fall 2010, leading to the development of the current 47-item version described below.

Overview of the TMACT

One goal in developing the TMACT was to create more explicit instructions to minimize rater subjectivity, thereby increasing reliability and validity of the scale itself. As a result, a robust instrument emerged. The TMACT protocol comprises three sections: (a) an Introduction (Part I) that orients the user to the fidelity evaluation process; (b) Itemized Data Collection Forms (Part II) to guide the user's interviews, observations, and ratings; and (c) Appendices, which include templates and data collection tools to facilitate the evaluation process.

The Introduction orients the user to the fidelity assessment process, including the intended primary use of fidelity data (i.e., to guide performance improvement efforts), as well as recommended fidelity raters' competencies and relationship to the assessed program. To ensure proficient use of this scale, the Introduction emphasizes evaluator competency in the underlying philosophy of ACT and other EBPs for this clinical population, as interpretation of unique situations and practices will undoubtedly occur and require judgment that goes beyond the detailed scoring guidelines provided in Part II. The Introduction also includes checklists, timelines, and considerations for preparing for, conducting, and finalizing a fidelity review.

Many of the forms in the Appendix provide specific templates for completing the fidelity review steps outlined in the Introduction. These include examples of a fidelity orientation letter, a fidelity review agenda, and a feedback report. The Appendix also contains several data collection forms to use throughout the evaluation, organized by data source. For example, a list of items that rely on chart data is incorporated into a worksheet.

Part II of the TMACT contains the heart of the instrument itself, where a 5-point behaviorally anchored scale follows each item's description and explicit rating guidelines. As with the DACTS, we begin by defining each item and including a rationale for its relative importance. We list data sources, making note of primary sources for the particular item, and provide guidance in how to synthesize data to determine whether item criteria have been met. Working from the DACTS protocol, we revised and expanded the interview questions to help elicit relevant information. Compared with the DACTS protocol, the TMACT protocol includes more explicit guidelines for scoring, such as data inclusion and exclusion criteria; tables and checklists that outline specific functions and criteria; case examples to help the rater judge whether criteria were met, either fully or partially; and formulae for ratings. An example of the protocol layout is provided in Table 1.

The TMACT includes 47 items making up the following six subscales: Operations and Structure, Core Team, Specialist Team, Core Practices, Evidence-Based Practices, and Person-Centered Planning and Practices. Table 2 lists all items according to subscale and a notation of its origin in relationship to the scale's predecessor, the DACTS. All items were changed in some manner, whether it was the definition and criteria used to assess the item or recalibrating the item anchors by redistributing data points. We removed several DACTS items (e.g., Continuity of Staffing) as we and other experts judged them to be less of a defining feature of ACT, albeit still recognized as important predictors and outcomes of strong programs. Furthermore, several original DACTS items were collapsed into one item (e.g., Responsibility for Hospital Admission and Hospital Discharge) or unbundled into several items (e.g., Full Responsibility for Services).

TMACT Subscales

Operations and Structure(OS) subscale. Twelve items contained in the OS subscale define the boundaries of the program, who and how many individuals the team is to serve, for how long, and the expectation of fostering continuity and ease of care across service and social systems. Also included in this subscale are more macro team processes that help keep the team on task, which include office-based program assistance and the daily team meeting. Instead of solely focusing on the daily team meeting attendance and frequency, as with the DACTS, we enhanced the fidelity scale by focusing the assessment on the quality of this essential meeting. Daily team meeting processes that are evaluated include the review of all consumers each day, documentation of relevant clinical information, and development of a daily schedule that is driven by the consumer's treatment plan, emerging needs

Table 1. TMACT Protocol—Item Example

OS1. Low Ratio of Consumers to Staff: Team maintains a low consumer-to-staff ratio of 10:1, including all direct service staff.

DATA SOURCES

Team Survey	See Item 2: Number of full-time ACT staff See Item 8: Number of consumers currently served

ITEM RESPONSE CODING

Inclusion Criteria	*ACT Staff:* Count all staff who provide direct services (substance abuse specialist, vocational specialist, and team leader) EXCEPT the psychiatric care provider. Part-time or temporary staff must work exclusively with the team for at least 16 hours/week and attend the daily team meeting at least 2 times per week. *Consumers:* In counting the current caseload, include all "active" or "enrolled" consumers. The caseload totals should include any consumer who has been formally admitted, even if it is as recent as the past week. This count should not exclude consumers currently enrolled on the team who are difficult to engage and have not had recent contact with the team. The definition of active status is determined by the team, but note that the count will affect other fidelity items, such as frequency of visits.
Exclusion Criteria	• Do not include psychiatric care provider in count. • Do not include administrative support staff, such as the program assistant or other managers assigned to provide administrative oversight to the team. • Do not count staff who are employed by the team but who have been on extended leave for *3 months* or more.
Formula	(# CONSUMERS CURRENTLY SERVED)/(# FTE STAFF)

	1	2	3	4	5
OS1. Low Ratio of Consumers to Staff	26 Consumers per team member or more	19-25	14-18	11-13	10 Consumers per team member or fewer

Note. TMACT = Tool for Measurement of Assertive Community Treatment; ACT = assertive community treatment.

(e.g., crises, dental or medical appointments), and proactive contacts to divert future crises. Also assessed is whether there is a mechanism to determine whether scheduled contacts were completed.

Core Team (CT) subscale. Seven items contained in the CT subscale assess staff positions that are critical to even the basic functioning of an ACT-like program. These individuals include the team leader, nursing staff, and a psychiatric care provider, which expands on DACTS criteria to also include qualified advanced practice nurse practitioners. As described below, staffing items assess not only whether a qualified individual is on the team but the services he or she provides to consumers and his or her role within the team.

Specialist Team (ST) subscale. Eight items contained in the ST subscale assess other staffing positions that contribute to the breadth of care provided to consumers. These include the substance abuse specialist, vocational specialist, and peer specialist. For each specialist, there is an expectation that the majority of services they provide incorporate elements of their area of specialty. This enhancement was in response to our and others' observations of a common problem with which many ACT teams struggle—scheduling specialists' time so that they may provide rehabilitative or therapeutic services that go beyond more maintenance tasks (e.g., medication monitoring, money allocation and management, and accessing groceries). Thus, in the TMACT, a warm, qualified body appointed to the position is not sufficient for high-fidelity standards; staff role items include standards that promote full integration within the team, service specialization, and application of EBPs. Specialists are expected to participate in the daily team meeting, contributing their perspective to clinical discussions, partake in assessment and treatment planning, and play an active role in cross-training and modeling skills for their fellow team members.

In developing TMACT items, we attempt to balance the importance of specialists delivering services within their area of training (including those integrated with engagement interventions), while also relying on the team, in total, to shoulder responsibility for all services, operating as "generalists." With cross-training among team members, the team may share more responsibility in delivering a high-quality specialty service. Initially, the specialist is thoughtfully scheduled to work with consumers who would benefit most from their area of specialty, but as the team becomes more competent in a given specialist area, the specialist can then comfortably step back and take a less prominent role as *the expert* on the team.

Core Practices (CP) subscale. Across eight items, core practices considered fundamental to the ACT model and to meeting consumers' basic needs are assessed. The

Table 2. The 47 Items Within the TMACT

Subscale Items		Item Origin
Operations and Structure (OS) subscale		
OS1	Low Ratio of Consumers to Staff	Revised DACTS item
OS2	Team Approach	Revised DACTS item
OS3	Daily Team Meeting (Frequency and Attendance)	Expanded DACTS item
OS4	Daily Team Meeting (Quality)	Added item
OS5	Program Size	Revised DACTS item
OS6	Priority Service Population	Expanded DACTS item
OS7	Active Recruitment	Expanded DACTS item
OS8	Gradual Admission Rate	Revised DACTS item
OS9	Graduation	Revised DACTS item
OS10	Retention Rate	Revised DACTS item
OS11	Coordination of Hospitalization	Collapsed two DACTS items
OS12	Dedicated Office-Based Program Assistance	Added item
Core Team (CT) subscale		
CT1	Team Leader on Team	Added item
CT2	Team Leader is Practicing Clinician	Revised DACTS item
CT3	Psychiatric Care Provider on Team	Revised DACTS item
CT4	Role of Psychiatric Care Provider (In Treatment)	Added item
CT5	Role of Psychiatric Care Provider (Within Team)	Added item
CT6	Nurses on Team	Revised DACTS item
CT7	Role of Nurses	Added item
Specialist Team (ST) subscale		
ST1	Substance Abuse Specialist on Team	Revised DACTS item
ST2	Role of Substance Abuse Specialist (In Treatment)	Added item
ST3	Role of Substance Abuse Specialist (Within Team)	Added item
ST4	Vocational Specialist on Team	Revised DACTS item
ST5	Role of Vocational Specialist (In Employment Services)	Added item
ST6	Role of Vocational Specialist (Within Team)	Added item
ST7	Peer Specialist on Team	Revised DACTS item
ST8	Role of Peer Specialist	Added item
Core Practices (CP) subscale		
CP1	Community-Based Services	Revised DACTS item
CP2	Assertive Engagement	Revised DACTS item
CP3	Intensity of Service	Revised DACTS item
CP4	Frequency of Contact	Revised DACTS item
CP5	Frequency of Contact With Natural Supports	Revised DACTS item
CP6	Responsibility for Crisis Services	Revised DACTS item
CP7	Full Responsibility for Psychiatric Services	Unbundled DACTS item
CP8	Full Responsibility for Rehabilitative Services	Unbundled DACTS item
Evidence-Based Practices (EP) subscale		
EP1	Full Responsibility for Dual Disorders Treatment	Unbundled DACTS item
EP2	Full Responsibility for Vocational Services	Unbundled DACTS item
EP3	Full Responsibility for Wellness Management Services	Added item
EP4	Integrated Dual Disorders Treatment (IDDT) Model	Revised DACTS item
EP5	Supported Employment Model	Added item
EP6	Engagement and Psychoeducation with Natural Supports	Added item
EP7	Empirically Supported Psychotherapy	Added item
EP8	Supportive Housing	Added item

(continued)

Table 2. (continued)

Subscale Items		Item Origin
Person-Centered Planning and Practices (PP) subscale		
PP1	Strengths Inform Treatment Plan	Added item
PP2	Person-Centered Planning	Added item
PP3	Interventions Target Broad Range of Life Goals	Added item
PP4	Consumer Self-Determination and Independence	Added item

Note. TMACT = Tool for Measurement of Assertive Community Treatment; ACT = assertive community treatment; DACTS = Dartmouth Assertive Community Treatment Scale. "Added" items are TMACT items that assess program features that were not assessed in the DACTS. "Revised" items are DACTS items that were modified significantly (e.g., the criteria used to make rating, or the calibration of the rating anchors) to create TMACT items, while maintaining the original concept for that item. "Collapsed" items are those TMACT items that assess a broader feature previously assessed with two or more narrowly defined DACTS items. "Unbundled DACTS items" are TMACT items that now assess a unidimensional concept, which were once included in a multidimensional DACTS item. "Expanded" items are those TMACT items that more fully assess a concept that was more minimally present in the DACTS.

essence of ACT core practices is how the team serves and supports a clinical population with complex needs and a lower level of functioning in the least restrictive environment. In addition to intense (i.e., of long duration) and frequent contact with consumers, the team aspires to develop relationships with consumers' natural support systems and is available 24/7 via crisis on call. In vivo contact has been a defining feature of ACT, where the team reaches out to consumers in their own environment, providing the vast majority of services in the community. The team is also equipped to work with poorly engaged consumers by using creative and nonintrusive outreach strategies and motivational interviewing. For those individuals where the stakes are too high in terms of safety or a significant loss of liberty because of high-risk behaviors, time-limited use of more restrictive interventions may be used. Finally, core practices also include psychiatric treatment and rehabilitation; a high-fidelity ACT program is highly competent in psychiatric rehabilitation and works individually with each consumer to help him or her develop daily living skills and work toward greater independence and self-sufficiency.

Evidence-Based Practices (EP) subscale. As ACT is typically viewed as a one-stop service provider, it is critical that not only an array of services be offered but that staff adopt best practices for a given service domain. A number of psychosocial EBPs have been identified for the clinical population targeted for ACT (Dixon et al., 2010), many of which are assessed across eight TMACT items within the EP subscale. These practices include integrated dual disorder treatment, supported employment, wellness management, supportive housing, cognitive–behavioral therapies, and family psychoeducation and support—each of which has its own extensive fidelity measure. The TMACT does not attempt to assess any of them comprehensively in just a few items, nor is it assumed that teams should necessarily reach the level of fidelity expected of clinical units uniquely dedicated to those practices. Instead, the TMACT assesses specialists' practices in light of clinical research evidence, and several items examine the entire team's adoption of a particular EBP philosophy and practice. To be practical, the assessment of team adoption of an EBP is cursory, and variability across team members is likely. Evaluators are, therefore, prompted to consider the balance of perspectives and judge whether individual criteria are met given the broader gestalt of the team. In the case of supported employment, teams are assessed across five criteria:

1. Values competitive work as a goal for all consumers
2. Believes that a consumer's expressed desire to work is the only eligibility criterion for supported employment services
3. On-the-job assessment is more valuable than extensive prevocational assessment
4. Placement is individualized and tailored to a consumer's preferences
5. Ongoing supports and job coaching is provided when needed and desired by the consumer

Not only is quality of services assessed but also penetration of these services. As a one-stop provider, consumers should not have to go outside the team for their service needs—if they are able to navigate a more disjointed care system, they may not need ACT. In the TMACT, several core and evidence-based services are assessed in terms of the extent to which the team is delivering the service, which is judged to be of at least adequate quality, to the those consumers who want and/or need the particular service. Thus, in the Full Responsibility items (see Table 2), the percentage of consumers determined to be receiving a particular service from the team is divided by the minimum percentage of consumers who likely need and/or want that service, where *need/want* is informed by published estimates, where available.[1]

Person-Centered Planning and Practices (PP) subscale. Four items assess the extent to which the team adopts person-centered assessment, treatment planning, and service delivery practices. Although elements of the recovery philosophy are an undercurrent of all TMACT items, this subscale takes a keen look at the team's approach to working collaboratively with consumers and focusing care on what consumers need and want. Although the PP subscale items tend to be more germane to any community mental health program, they are emphasized here for several reasons: ACT teams have a potential for assuming a more paternalistic role with consumers given their heightened supervisory role, and multiple staff working together to address consumers' needs arguably requires more sophisticated treatment planning and execution.

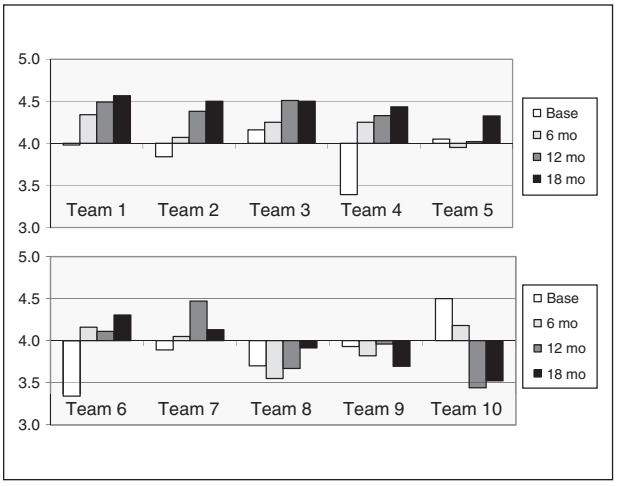

Figure 1. TMACT total scores for 10 Washington ACT teams at the 4 assessment points

TMACT Pilot in One State

Method

Participants. Ten ACT teams in Washington State were administered the TMACT as part of their state-contracted fidelity assessment requirements. These teams consisted of (a) six designated as "full" teams, serving 90 to 100 consumers and employing an average of 11 full-time staff and one psychiatric care provider for every 50 consumers and (b) four designated as "half" teams serving 42 to 50 consumers and employing an average of 7 full-time staff and one psychiatric care provider for every 50 consumers. Seven of these teams were implemented in Western Washington in July 2007, while three were implemented in Eastern Washington in October 2007. Data were collected as part of ongoing operational state agency oversight and were associated with organizational units—the teams, not individual persons, either staff or consumers—so that human subjects as such were not involved at this stage.

Materials and procedure. Fidelity reviews were conducted at baseline (4-6 months post-implementation), 6, 12, and 18 months using the TMACT.[2] Forms were created to also capture data relevant for only the DACTS, so that evaluators could simultaneously rate teams on both fidelity measures, for comparison. Two trained fidelity reviewers—one from the University of Washington and one from the Washington State Division of Behavioral Health and Recovery—conducted each review over the course of 1½ to 2 days. Typical data sources included the following: (a) team self-report on a survey regarding team staffing, credentials, and other objective team information (e.g., number of consumers served, number of discharges); (b) team self-report of consumer-level data on services received and hospitalizations over the review period; (c) observation of the daily team meeting; (d) observation of a treatment planning meeting; (e) review of 10 randomly selected charts; (f) interviews with the team leader, psychiatric care provider, nurses, clinicians, peer specialist, substance abuse specialist, and vocational specialist; and (g) interviews with 3 to 5 consumers. Depending on time and logistical constraints with each review, fidelity reviewers also shadowed staff during home and community visits, conducting informal interviews with staff and consumers during this time.

Results

Figure 1 shows total mean TMACT scores for each of the 10 teams at the 4 assessment points. Teams' results are numbered and shown on two rows in descending order of 18-month follow-up TMACT scores. Although the scale goes from 1 to 5, no team received a score below 3.0, so only the 3 to 5 range is shown. The *x*-axis is placed at 4.0 both for improved clarity and because mean scores above 4.0 indicate substantial adherence to the components of the model measured.

Six of the 10 teams showed relatively consistent progress over the time period measured and concluded with scores at the 18-month point well above 4.0. Of the four other teams with significantly lower total scores (3.82 ± 0.26 vs. 4.44 ± 0.01; $t[3.63] = 4.49$, $p = .014$), two had a more variable course but finished close to or above 4.0. The other two teams showed a generally declining course, concluding much lower, even though one had started with a relatively high score. The patterns of observed TMACT scores are consistent with informal observations and anecdotal reports. All the teams doing less well had a greater degree of problem with recruitment and retention of key staff, including psychiatric care providers, nurses, and the team leader, thus creating disruptions in continuity of care and expertise and necessarily compromising multiple aspects of high-fidelity performance. These teams

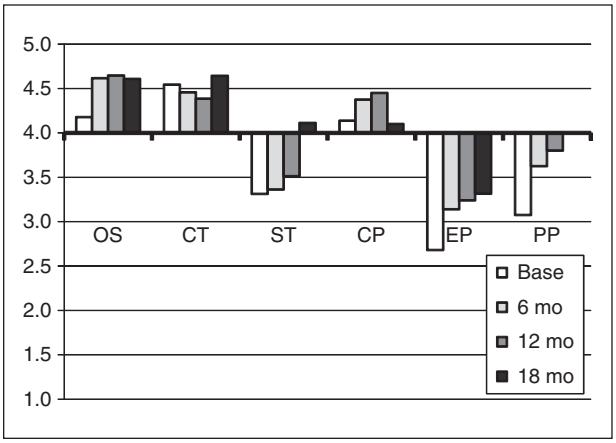

Figure 2. TMACT subscales: mean scores for 10 Washington teams at the 4 assessment points
Note. Subscales: OS = Operations and Structure; CT = Core Team; ST = Specialist Team; CP = Core Practices; EP = Evidence-Based Practices; PP = Person-Centered Planning and Practices.

Table 3. Comparison of Mean DACTS and TMACT Scores at 4 Points in Time for 10 ACT Teams in Washington State

	DACTS		TMACT			Paired Comparisons		
Time	Mean	*SD*	Mean	*SD*	Difference	*T*	*df*	*p*
Baseline	4.36	0.19	3.88	0.11	0.48	5.34	9	.000
6 months	4.27	0.07	4.06	0.08	0.21	4.85	9	.001
12 months	4.35	0.08	4.14	0.12	0.21	4.54	9	.001
18 months	4.36	0.08	4.19	0.12	0.17	2.08	9	.067

Note. TMACT = Tool for Measurement of Assertive Community Treatment; *SD* = standard deviation; ACT = assertive community treatment; *df* = degrees of freedom; DACTS = Dartmouth Assertive Community Treatment Scale.

also encountered other systemic or organizational issues (e.g., changing from a three-agency team to a one-agency team). However, Figure 1 also suggests that the overall level of performance of the teams was relatively high, especially for new teams, likely a function of the strong investment in training and support provided.

Figure 2 shows scores on TMACT subscales across all 10 teams at the 4 assessment points. As expected, scores for the three scales focusing on more traditional ACT criteria—OS, CT, and CP—were higher than those for the three scales more explicitly oriented to recovery and EBPs—ST, EP, and PP (4.45 ± 0.26 vs. 3.81 ± 0.55; $t[9] = 5.26$, $p = .001$). Recovery/EBP program dimensions are more difficult to implement—especially ensuring whole-team involvement and sufficient penetration with EBPs—and take longer to implement, but these teams overall demonstrated consistent progress with them over time. The observed substantial drop in scores for CP is consistent with the challenges a minority of the teams found in maintaining service quality in the context of turnover of key staff.

Table 3 shows a comparison of overall TMACT and DACTS scores across all 10 teams at each time point. The high DACTS scores confirm that, compared with samples elsewhere, these teams as a group exhibited quite high fidelity as indicated by this instrument (Salyers et al., 2003). DACTS scores remained essentially unchanged and were significantly higher than TMACT scores over most of the observation period. In contrast, TMACT scores rose steadily during the observation period; by the 18-month follow-up point, the difference between the two scales was of only marginal significance.

The differences in patterns of these two scales, both overall and with respect to TMACT subscales, provide preliminary evidence that the TMACT is performing as intended. First, it sets a higher bar for performance, especially through assessing the recovery/EBP dimensions now deemed essential to high-fidelity ACT. The higher scores on the DACTS are consistent with higher scores on OS, CT, and CP subscales of the TMACT, with which the DACTS shares most of its content. Omission of the additional content provided by other subscales on the TMACT would provide a misleading estimate of performance. Second, the TMACT scores indicate that the measure is sensitive to change over time—consistent with expected improvement through training and maturation of teams—whereas the DACTS shows no such change. Third, improvement in sensitivity appears to extend to detection of differences between teams. The four teams that, as a group, had significantly lower 18-month TMACT scores than the other six at the $p = .001$ level showed only marginally lower DACTS scores (4.18 ± 0.24 vs. 4.48 ± 0.19; $t[8] = 2.24$, $p = .056$).

Discussion and Next Steps

Through an extensive, iterative process of development, a more robust ACT fidelity tool has emerged. As indicated in the results described above, the TMACT sets a higher bar for ACT program performance and is more sensitive to change than the DACTS. Furthermore, variations across subscale ratings match expectations of challenges in implementing ACT components, including the challenges in fully implementing specialist roles and associated EBPs, as well as integrating person-centered, recovery-oriented practices.

The value of having an ACT fidelity measure with greater specificity is significant for the administrative and clinical fields. Unlike the DACTS, on which teams clearly lacking essential elements were nonetheless categorized as "high fidelity," the TMACT distinguishes among ACT programs at different levels of functioning

and quality. More precise information about performance will be helpful in targeting training, consultation, supervision, and general performance improvement efforts at both the program and statewide levels.

Furthermore, a stronger measure of ACT has the potential for advancing the field of research on ACT. Although earning the designation as an "evidence-based practice" after decades of research, ACT remains one of the most costly community mental health programs, and questions remain regarding exactly which features and what dose are most critical to desired outcomes for various groups of consumers. At the same time, there is a limited range of outcomes for which the evidence of ACT effectiveness has been unequivocal—specifically, reduction in psychiatric hospital use, increase in housing stability, and engagement in and satisfaction with treatment (Bond, Drake, Mueser, & Latimer, 2001). As the field has come to embrace the concepts embodied in the New Freedom Commission on Mental Health (2003) report, which placed recovery at the center of the mental health enterprise and underscored the need to provide evidence-based care, the array of potentially desired outcomes has broadened, and the implicit mandate for ACT teams has expanded accordingly. Inclusion of a range of specialists on staff in ACT teams is already recognized as imperative (Bond et al., 2001); as the predominantly exclusive provider of care for their consumers, teams would also be expected to ensure that these specialists provide to the degree possible the evidence-based treatments their consumers need. Although the kind of commitment that has been made in Washington State to emulating other EBPs within ACT is not universally accepted, our informants within the ACT provider community consistently supported this direction for the model. Inclusion of items in a comprehensive fidelity measure that tap not just the presence of specialists on the team but also the degree to which their activities are informed by evidence for effective practice in their areas will improve prospects for future research on ACT effectiveness in respective additional recovery-focused domains.

With the passing of mental health parity, we can expect that the demand for mental health services may increase, leading administrators and funders to pay even greater attention to what works for whom at what cost. In a resource-constrained mental health system, providers often struggle to implement EBPs, and recent research has shown the importance of boosting facilitators and removing barriers at multiple levels of the system, from the state mental health authority down to staff (Mancini et al., 2009). In setting the bar higher for high-fidelity ACT implementation, it will be important that administrators and providers foster an environment that increases the probability for successful implementation of a high-fidelity ACT team. The TMACT may offer a useful tool to support those efforts as it not only provides a more current conceptualization of ACT but is in itself a robust evaluation kit designed to assist evaluators in providing thoughtful quality improvement feedback to guide performance enhancements.

Since its initial development, the TMACT has drawn increasing attention by the ACT community, and several steps will be taken to ensure appropriate dissemination. At the time of this writing, the TMACT has been adopted and piloted in five U.S. states, Japan, and Norway, with some piloting by teams in three additional U.S. states. After several years of modifying the TMACT based on feedback from these initial pilot efforts, Version 1.0 of the TMACT was finalized in December 2010. The authors plan to continue to use and provide additional training in the locations where the TMACT has already been adopted while responding to current and ongoing requests for training and consultation from several additional states. Training materials will be refined with a focus on further enhancing validity and reliability of the measurement process. The authors will also continue to develop and pilot-test new approaches to supporting initial training and sustainability in application of the tool, for example, by training ACT team members in peer-based fidelity assessment of other teams (Teague & Monroe-DeVita, in press) or training a state-based training and technical assistance center to conduct ongoing training of fidelity reviewers (Salyers, McKasson, Bond, & McGrew, 2007).

At the same time, it is apparent that, as currently configured, a TMACT assessment represents a substantial investment in time and effort, regardless of how the cost is allocated. Most of the process relies on in-person, on-site review by two experienced fidelity evaluators, and the methods include only limited use of technological approaches to minimize burden. State mental health authorities are understandably concerned with the high cost of maintaining high-fidelity programs. Fidelity reviews are essential to ensure programs' investment value, but when conducted in this way represent a cost in their own right; we estimate the cost of a full review to be approximately 0.2% of the annual cost of an ACT team. However, in view of the high cost of ACT teams and of the potential value of accurate assessment and consequent gains in performance, the fidelity investment is justifiable, particularly for new teams and at least periodically for more mature teams; the marginal savings in acute care costs forgone as a result of improved performance could alone offset the costs of targeted fidelity assessments.

Nonetheless, future plans include exploration of distance technologies and methods to narrow the scope of review for follow-up with teams that have previously demonstrated good fidelity. McGrew, Salyers and colleagues (personal communication, September 24, 2009) have seen promising results with DACTS assessments

using distance technologies to reduce effort, and newer Internet-based communication methods could further reduce burden. Further research with the TMACT may indicate the feasibility of trimming items from the measure. It will be important in any of these processes not to induce selective attention to fidelity criteria and thereby undermine the validity of assessments, as discussed earlier. An alternative strategy such as using a subset of both targeted and randomly selected items in follow-up reviews in combination with distant but directly interactive technologies could significantly reduce burden while minimizing potential threat to model integrity.

Future research should address several questions related to the TMACT. First, an assessment of the psychometric properties of the tool is needed with a larger sample of programs. While the measure was being developed, no formal test of interrater reliability was conducted. Experience during development indicated that initial item score differences found within rater teams quickly become consensus ratings through shared review of the evidence; a formal test on a larger scale using the final measure will be necessary to evaluate whether this holds up more generally. Second, the important question of relationship of fidelity as measured by the TMACT to desired outcomes should also be evaluated on a larger and more diverse set of teams. As noted, ACT—high-fidelity ACT in particular—has been found to have an important limited set of well-established outcomes. But previously there has been only limited assessment of ACT ingredients in relation to these and other outcomes. The more comprehensive TMACT may provide a useful tool to evaluate not only overall effectiveness over the broad range of outcomes now expected of services for this population but also the extent to which specific ingredients are critical for specific outcomes.

Research of this type is likely to lead to eventual refinements in the measure. Though some stability is important for the purposes of longitudinal comparisons, we have seen that the ACT model evolves, and it will also be necessary to refine the TMACT over time to keep up. In the meantime, it appears that the changes incorporated in the TMACT have helped move toward better discrimination among low-, medium-, and high-fidelity ACT teams while also providing more specific feedback to guide ongoing performance improvement efforts and a basis for advancing knowledge about the critical ingredients of this important service model.

Acknowledgments

The authors wish to thank the numerous contributors to the development of the TMACT, particularly the late Deborah Allness, MSSW, who early on envisioned the need for such an enhancement.

Author Roles

Maria Monroe-DeVita, lead writing, editing, and submission of the manuscript; Gregory Teague, writing and editing; and Lorna Moser, writing and editing.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research and/or authorship of this article: This project was partially funded by the Washington State Division of Behavioral Health and Recovery.

Notes

1. In estimating the percentages of consumers who need and/or want Core and EBP services, as assessed in all five Full Responsibility items, we referred to published data and considered evidence for the clinical population targeted by ACT, effect of consumer choice, and measurement error. Given these considerations, we selected more conservative estimates to use only in default if a team reports a lower rate of service need/want for their caseload (i.e., we use the teams' estimates if they exceed the more conservative base rates). We chose a 40% base rate for estimating need/want for substance abuse services (Dixon, Haas, Weiden, Sweeney, & Frances, 1991; Drake & Wallach, 1989; Regier et al., 1990). We also chose 40% for need/want for vocational services, where estimates tend to range from 50% to 70% (McQuilken et al., 2003; Mueser, Salyers, & Mueser, 2001; Rogers, Walsh, Masotta, Danley, & Smith, 1991). However, little research has been reported on ACT consumers specifically, and we chose again to select a more conservative base rate. Although we believe that most or nearly all ACT consumers likely need/want psychiatric, rehabilitative, and wellness management (defined broadly) services, we use the base rate of 90%.
2. There were several revisions to the TMACT during this time frame, the most significant of which included the addition of two new items: Engagement and Psychoeducation with Natural Supports and Empirically Supported Psychotherapy. During this time frame, the Supportive Housing item had not yet been developed.

References

Allness, D. J. (2003). *National program standards for ACT teams* [Online guidelines]. Retrieved from http://www.nami.org/Template.cfm?Section=ACT-TA_Center

Allness, D. J., & Knoedler, W. H. (1998). *The PACT model of community-based treatment for persons with severe and persistent mental illnesses: A manual for PACT start-up.* Arlington, VA: National Alliance for the Mentally Ill.

Becker, D. R., Smith, J., Tanzman, B., Drake, R. E., & Tremblay, T. (2001). Fidelity of supported employment programs and employment outcomes. *Psychiatric Services*, *52*, 834-836.

Bjorklund, R., Monroe-DeVita, M., Reed, D., Toulon, A., & Morse, G. (2009). Washington State's initiative to disseminate and implement high-fidelity ACT teams. *Psychiatric Services*, *60*, 24-27.

Bond, G. R., Drake, R. E., Mueser, K. T., & Latimer, E. (2001). Assertive community treatment for people with severe mental illness. *Disease Management & Health Outcomes*, *9*, 141-159.

Bond, G. R., Evans, L., Salyers, M. P., Williams, J., & Kim, H.-W. (2000). Measurement of fidelity in psychiatric rehabilitation research. *Mental Health Services Research*, *2*, 75-87.

Bond, G. R., Salyers, M. P., Rollins, A. L., Rapp, C. A., & Zipple, A. M. (2004). How evidence-based practices contribute to community integration. *Community Mental Health Journal*, *40*, 569-588.

Brekke, J. S. (1988). What do we really know about community support programs? Strategies for better monitoring. *Hospital & Community Psychiatry*, *39*, 946-952.

Brekke, J. S., & Test, M. A. (1992). A model for measuring the implementation of community support programs: Results from three sites. *Community Mental Health Journal*, *28*, 227-247.

Dixon, L. B., Dickerson, F., Bellack, A. S., Bennett, M., Dickinson, D., Goldberg, R. W., . . . Kreyenbuhl, J. (2010). The 2009 schizophrenia PORT psychosocial treatment recommendations and summary statements. *Schizophrenia Bulletin*, *36*, 48-70.

Donabedian A. 1988. The quality of care. How can it be assessed? *Journal of the American Medical Association 260*(12):1743-1748.

Dixon, L., Haas, G. L., Weiden, P., Sweeney, J. K., & Frances, A. J. (1991). Drug abuse in schizophrenic patients: Clinical correlates and reasons for use. *American Journal of Psychiatry*, *148*, 224-230.

Drake, R. E., McHugo, G. J., Clark, R. E., Teague, G. B., Xie, H., Miles, K., . . . Ackerson, M. S. W. (1998). Assertive community treatment for patients with co-occurring severe mental illness and substance use disorder: A clinical trial. *American Journal of Orthopsychiatry*, *68*, 201-215.

Drake, R. E., & Wallach, M. A. (1989). Substance abuse among the chronic mentally ill. *Hospital & Community Psychiatry*, *40*, 1041-1046.

Latimer, E. (1999). Economic impacts of assertive community treatment: A review of the literature. *Canadian Journal of Psychiatry, 44*, 443-454.

Mancini, A. D., Moser, L. L., Whitley, R., McHugo, G. J., Bond, G. R., Finnerty, M. T., & Burns, B. J. (2009). Assertive community treatment: Facilitators and barriers to implementation in routine mental health settings. *Psychiatric Services*, *60*, 189-195.

McGrew, J. H., Bond, G. R., Dietzen, L., & Salyers, M. (1994). Measuring the fidelity of implementation of a mental health program model. *Journal of Consulting and Clinical Psychology*, *62*, 670-678.

McHugo, G. J., Drake, R. E., Teague, G. B., & Xie, H. (1999). Fidelity to assertive community treatment and client outcomes in the New Hampshire Dual Disorders Study. *Psychiatric Services*, *50*, 818-824.

McHugo, G. J., Hargreaves, W., Drake, R. E., Clark, R. E., Xie, H., Bond, G. R., & Burns, B. J. (1998). Methodological issues in assertive community treatment studies. *American Journal of Orthopsychiatry*, *68*, 246-260.

McQuilken, M., Zahniser, J. H., Novak, J., Starks, R. D., Olmos, A., & Bond, G. R. (2003). The work project survey: Consumer perspectives on work. *Journal of Vocational Rehabilitation*, 18, 59-68.

Moncher, F. J., & Prinz, R. J. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review*, *11*, 247-266.

Mowbray, C. T., Holder, M., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *American Journal of Evaluation*, *24*, 315-340.

Mueser, K. T., Salyers, M. P., & Mueser, P. R. (2001). A prospective analysis of work in schizophrenia. *Schizophrenia Bulletin*, *27*, 281-296.

New Freedom Commission on Mental Health. (2003, July 22). *Achieving the promise: Transforming mental health care in America. Final report* (DHHS Publ. No. SMA-03-3832). Rockville, MD: Author.

Phillips, S. D., Burns, B. J., Edgar, E. R., Mueser, K. T., Linkins, K. W., Rosenheck, R. A., . . . McDonel Herr, E. C. (2001). Moving assertive community treatment into standard practice. *Psychiatric Services*, *52*, 771-779.

Regier, D. A., Farmer, M. E., Rae, D. S., Locke, B. Z., Keith, S. J., Judd, L. L., & Goodwin, F. K. (1990). Comorbidity of mental disorders with alcohol and other drug abuse. Results from the Epidemiologic Catchment Area (ECA) Study. *Journal of the American Medical Association*, *264*, 2511-2518.

Rogers, E. S., Walsh, D., Masotta, L., Danley, K., & Smith, K. (1991). *Massachusetts survey of client preferences for community support services: Final report* (Submitted to the Department of Mental Health). Boston, MA: Boston University, Center for Psychiatric Rehabilitation.

Salyers, M. P., Bond, G. R., Teague, G. B., Cox, J. F., Smith, M. E., Hicks, M. L., . . . Koop, J. (2003). Is it ACT yet? Real-world examples of evaluating the degree of implementation for assertive community treatment. *Journal of Behavioral Health Services & Research*, *30*, 304-320.

Salyers, M. P., McKasson, M., Bond, G. R., & McGrew, J. H. (2007). The role of technical assistance centers in implementing evidence-based practices: Lessons learned. *American Journal of Psychiatric Rehabilitation*, *10*, 85-101.

Salyers, M. P., &Tsemberis, S. J. (2007). ACT and recovery: Integrating evidence-based practice and recovery orientation on assertive community treatment teams. *Community Mental Health Journal*, *43*, 619-641.

Stein, L. I., & Test, M. A. (1980). Alternative to mental hospital treatment: I. Conceptual model, treatment program,

and clinical evaluation. *Archives of General Psychiatry*, *37*, 392-397.

Substance Abuse and Mental Health Services Administration. (2010). Publications ordering: Evidence-based practices. retrieved September 15, 2010, from http://store.samhsa.gov/facet/Professional-Research-Topics/term/Evidence-Based-Practices

Teague, G. B., Bond, G. R., & Drake, R. E. (1998). Program fidelity in assertive community treatment: Development and use of a measure. *American Journal of Orthopsychiatry*, *68*, 216-232.

Teague, G. B., Drake, R. E., & Ackerson, T. (1995). Evaluating use of continuous treatment teams for persons with mental illness and substance abuse. *Psychiatric Services*, *46*, 689-695.

Teague, G. B., & Monroe-DeVita, M. (in press). Not by outcomes alone: Using peer evaluation to ensure fidelity to evidence-based assertive community treatment (ACT) practice. In J. L. Magnabosco & R. W. Manderscheid (Eds.), *Outcomes measurement in the human services: Cross-cutting issues and methods* (2nd ed.). Washington, DC: National Association of Social Workers Press.

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