

Introduction to the Evidence-Based Practice Fidelity Scales

What is fidelity?

Fidelity refers to the degree of implementation of an evidence-based practice (EBP), and a scale that measures fidelity is called a *fidelity scale* (Bond, Evans, Salyers, Williams, & Kim, 2000). Such scales have been developed for each of the six EBPs included in the Implementing EBP Project: assertive community treatment (ACT), supported employment (SE), illness management and recovery (IMR), family psychoeducation, integrated treatment for dual disorders treatment (IDDT), and medication management according to protocol (MedMAP). Each scale consists of approximately 15 – 30 items assessing the critical ingredients of the EBP, based on the underlying principles and methods of the EBP. The scale items provide concrete indications that the practice is being implemented as intended. For example, one item on the Supported Employment Fidelity Scale concerns *Rapid Job Search*. This item is rated as fully implemented if the consumers in a program average one month or less between admission to the supported employment program and their first job interview.

Why measure fidelity?

Several assumptions underlie the use of fidelity scales. First, a fidelity scale should adequately sample the critical ingredients of the EBP in order to differentiate between programs that follow the practice and those that do not. This is sometimes referred to as *discriminant validity*. Second, fidelity scales should be sensitive enough to detect progress in the development of a program from the start-up phase to its mature development. Third, high-fidelity programs are expected to have greater effectiveness than low-fidelity programs in achieving desired outcomes. This is referred to as *predictive validity*.

One key use of fidelity scales is for monitoring programs over the course of their development (and even after they are fully established). Programs that are not monitored may erode over time in their fidelity of implementation. Routine use of fidelity scales provides an objective, structured way to give feedback about program development. Experience also suggests that this is an excellent method to diagnose program weaknesses, while helping to clarify program strengths. Fidelity scales also provide a comparative framework for evaluating statewide implementation of an EBP and allow the evaluator to identify statewide trends and outliers. The strategic use of repeated evaluations of programs using fidelity scales, either for an individual program or on a statewide level, is based on the general principle that whatever is attended to is more likely to be improved.

How are fidelity scales used?

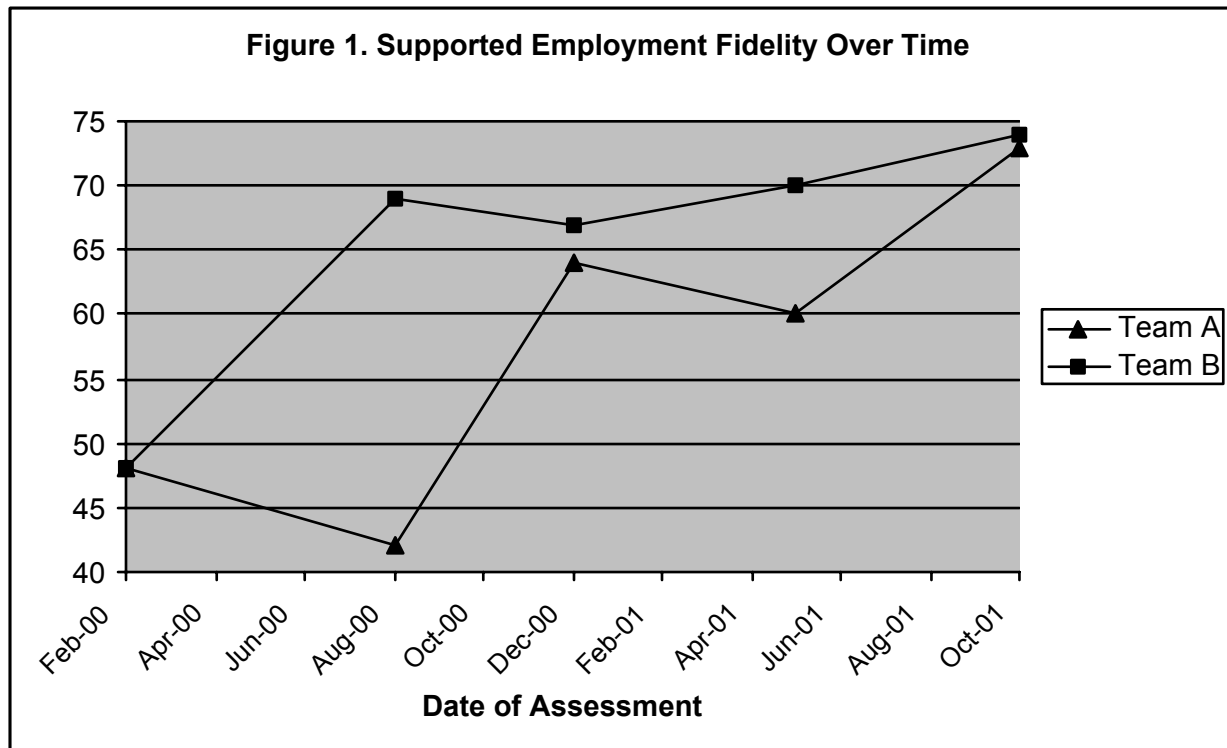
In the Implementing EBP Project we have developed fidelity scales consisting of simple-to-understand, face-valid items that are rated on a 5-point response format, ranging from 1 = not implemented to 5 = fully implemented, with intermediate numbers representing progressively greater degrees of implementation. The response alternatives are behaviorally anchored, that is, they identify concrete measurable elements of the practice. Our experience is that consensus ratings derived from independent evaluators using multiple sources of information make the most valid ratings. Typical sources of information include interviews with staff, observation of team meetings, review of charts, and observation of interventions. A day-long site visit is the optimal method for acquiring this information. Interviewers should be familiar with the EBP being rated.

Although we recommend outside raters, program managers can also use fidelity scales to conduct self-ratings. The validity of self-ratings (or any ratings, for that matter) depends on the knowledge of the person making the ratings, access to accurate information pertaining to the

ratings, and the objectivity of the ratings. We encourage the use of self-ratings, with appropriate caveats regarding potential biases that can be introduced by raters who are invested in seeing a program “look good” or who do not fully understand the principles of the EBP. In addition to the scales developed for independent evaluators and program managers, companion fidelity measures use by consumers and family members are under development for some EBPs.

Graphing fidelity ratings

We recommend that programs implementing an EBP graph their fidelity ratings over time, using their total fidelity score. By graphing this score, a program can see its change over time. An example is given in Figure 1. When the program shows greater fidelity over time, this will be displayed by the graph and serve to reinforce their efforts. Another feature of graphing fidelity is to examine the cut-off score for “full implementation.” A program can use this score as a target and can gauge how they measure up.



Note: 66 - 75 = good implementation 56 - 65 = fair implementation 55 & below = not SE

Status of the EBP Fidelity Scales

As noted above, fidelity scales have now been developed for each of the six EBPs. Two of the scales (for assertive community treatment and supported employment) were developed prior to the start of the Implementing EBP Project, and draft instruments for the remaining four were developed between January 2000 – November 2002. A summary of field experience with actual application of these scales is given in Table 1.

Table 1. Status of EBP Fidelity Scales (11/18/02)

EBP	Scale Developed?	Piloted	Discriminant Validity?	Predictive Validity?	Broad Use?
ACT	1998		Yes	Yes	Yes
SE	1997		Yes	Yes	Yes
IMR	2002	In progress			
Family	2002	In progress			
IDDT	2002	In progress			
Medications	2002	Planned			

Regarding reliability of ratings using the fidelity scales, our experience is that agreement between fidelity assessors is generally very good. Internal consistency of the scales has been satisfactory when a range of programs have been sampled. Little information is available at this time regarding test-retest reliability. Information from different data sources does not always agree, however (Winter & Calsyn, 2000). However, by incorporating multiples sources of data, we believe raters will be able to better approximate “true” scores. Through the development of protocols for administering and coding the scales, we hope to achieve better reliability than has been found in some early studies.

The two scales with longer histories have been subject to psychometric studies. ACT is measured by the Dartmouth ACT Fidelity Scale (DACTS), which was developed and described by Teague, Bond, and Drake (1998). As shown in Table 1, the DACTS has demonstrated discriminant and predictive validity and, in fact, has been widely adopted by many state and local agencies throughout the United States and internationally. The DACTS has been found to differentiate between established ACT teams, as monitored and trained by ACT trainers, and other types of intensive case management and brokered case management (Teague et al, 1998). Regarding predictive validity, several studies using precursors to the DACTS have found strong correlations between ACT fidelity and client outcomes (L atimer, 1999; McGrew, Bond, Dietzen, & Salyers, 1994; McHugo, Drake, Teague, & Xie, 1999). One study using the DACTS found a .49 correlation between the DACTS and reduction in hospital use (which was not statistically significant, perhaps due to small sample size) (Bond & Salyers, submitted).

The Supported Employment Fidelity Scale (formerly known as the IPS Fidelity Scale) was developed and described by Bond, Becker, Drake, and Vogler (1997). The scale differentiates between well-established programs following the Individual Placement and Support model of supported employment and other types of vocational programs, as well as other programs labeled as supported employment (Bond et al, 2001). One study found a strong correlation (.76) between fidelity, as rated by the Supported Employment Fidelity Scale and competitive employment rates (Becker, Smith, Tanzman, Drake, & Tremblay, 2001).

For the remaining four EBPs, a group of researchers at Indiana University Purdue University Indianapolis (Gary Bond, Kikuko Campbell, Jane Williams, Michelle Salyers, Angie Rollins, and Natalie DeLuca) have worked with the developers of each of the Implementation Resource Kits to create fidelity scales: IMR (Kim Mueser and Susan Gingerich), Family

Psychoeducation (Bill McFarlane and Lisa Dixon), IDDT (Bob Drake), and MedMAP (Alex Miller). These scales have undergone numerous drafts and review by many different groups, including the implementation monitors and trainers who are using these scales in Phase II of the Implementing EBP Project. As of November, 2002, these four scales have not been formally piloted, but they have been refined and have undergone some field testing. An earlier version of the IDDT fidelity Scale has been extensively piloted in Ohio, with promising results regarding its utility.

In addition to the EBP-specific fidelity scales, we developed the General Organizational Index, a scale that measures fidelity to good clinical practice standards on a broader level. Bob Drake and Charlie Rapp were the primary developers of this measure. This scale helps quantify such aspects of practice as consumer choice, training, supervision, and monitoring of fidelity and outcomes. Like the four other newly-developed scales, this scale has undergone multiple revisions based upon feedback from a variety of sources (including the Phase II implementation monitors and trainers), and it continues to be refined as part of the EBP project process.

References

- 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.
- Bond, G. R., Becker, D. R., Drake, R. E., & Vogler, K. M. (1997). A fidelity scale for the Individual Placement and Support model of supported employment. Rehabilitation Counseling Bulletin, *40*, 265-284.
- Bond, G. R., Evans, L., Salyers, M. P., Williams, J., & Kim, H. K. (2000). Measurement of fidelity in psychiatric rehabilitation. Mental Health Services Research, *2*(2), 75-87.
- Bond, G. R., & Salyers, M. P. (submitted). Prediction of outcome from the Dartmouth ACT Fidelity Scale.
- Bond, G. R., Vogler, K. M., Resnick, S. G., Evans, L. J., Drake, R. E., & Becker, D. R. (2001). Dimensions of supported employment: Factor structure of the IPS Fidelity Scale. Journal of Mental Health, *10*, 383-393.
- Latimer, E. (1999). Economic impacts of assertive community treatment: A review of the literature. Canadian Journal of Psychiatry, *44*, 443-454.
- McGrew, J. H., Bond, G. R., Dietzen, L. L., & Salyers, M. P. (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). The relationship between model fidelity and client outcomes in the New Hampshire Dual Disorders Study. Psychiatric Services, *50*, 818-824.
- 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.
- Winter, J. P., & Calsyn, R. J. (2000). The Dartmouth ACT Scale: A generalizability study. Evaluation Review, *24*, 319-338.