

Proceedings Writeup

“Is it wraparound yet?”

Determining fidelity standards for the Wraparound Fidelity Index

Author Information:

Eric J. Bruns, Ph.D.

Assistant Professor

Department of Psychiatry and Behavioral Sciences

Division of Public Behavioral Health and Justice Policy

University of Washington School of Medicine

146 N. Canal St., Suite 100

Seattle, WA 98103

206-685-2477 (ph)

206-685-3430 (fax)

ebruns@u.washington.edu

Kristen Leverentz-Brady, M.A.

Graduate Assistant and Project Coordinator, Wraparound Evaluation and Research Team

Department of Psychology

University of Vermont

John Dewey Hall

Burlington, VT 05405

802-656-5002

nabbie69@yahoo.com

Jesse C. Suter, M.A.

Graduate Assistant

Department of Psychology

University of Vermont

John Dewey Hall

Burlington, VT 05405

603-542-5449

jsuter@uvm.edu

Acknowledgment: This study was supported in part by ORC Macro, Inc., and the U.S. Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, Child, Adolescent, and Family Branch.

“Is it wraparound yet?”

Determining fidelity standards for the Wraparound Fidelity Index

Eric J. Bruns, Kristen M. Leverentz-Brady, & Jesse C. Suter

Introduction

In recent years, the wraparound process for planning and implementing services and supports for children and youth with intensive needs has been cited as a promising service delivery option in major reviews (e.g., Burns, Hoagwood, & Maultsby, 1998) and Surgeon General's reports on both mental health and youth violence (USPHS, 1999, 2001). Meanwhile, significant efforts have been undertaken to better specify the wraparound model, including descriptions of specific provider and team member activities (Bruns, Walker, et al., 2004), refinement of the wraparound principles (Walker, Bruns et al., 2004), and necessary system and program supports for the model (Walker, Koroloff, & Schutte, 2003). Finally, research is beginning to demonstrate linkages between adherence to the wraparound principles and outcomes for youth (e.g., Bruns, Rast et al., in press; Bruns, Suter et al., 2005). Such developments have helped the wraparound process move from being perceived as merely a philosophy to a specified but flexible practice model with potential for positive impact.

At the same time, treatment fidelity, the degree to which a program is implemented as intended, has emerged as a major issue in behavioral health service delivery. As described by Salyers et al. (2003), fidelity assessment is "the natural union of scientific and practical needs of documenting and describing service provision" (p.305). With the human services field becoming increasingly technocratic and focused on implementation of evidence-based practices, fidelity assessment is increasingly employed by programs or trainers who need to conduct quality assurance activities; agencies who need to make funding and accreditation decisions; and researchers who need to interpret study results and comment on program quality.

With the wraparound process gaining acceptance as a specified program model, and agencies increasingly interested in using data to guide policy, funding, and certification decisions, there is a serious need for methods to determine when wraparound implementation in a program or community is "good enough." Not surprisingly, to date, such practical needs have outstripped the science of fidelity measurement for this practice model. Though several fidelity measures have been developed for the wraparound process, and their reliability and validity have been established (see, e.g., Bruns, Burchard et al., 2005), an empirical approach to determining what scores represent faithful implementation has not previously been attempted.

The current study aimed to "bootstrap" fidelity benchmarks to help programs or communities interpret scores derived from the Wraparound Fidelity Index, version 3 (WFI-3). As described by Salyers et al (2003), there are two main methods for interpreting assessment results. The first is to take a *norm-referenced* approach, in which a score for an individual (or program site) is compared to a large group of assessed individuals (or sites) to see how they compare. The second method would be to use a *criterion-referenced* approach, whereby a score is compared to an external standard, such as an one that is related to prediction of performance (e.g., a score that has been found to be associated with positive client outcomes or better service delivery).

In the current study, we used a combination of norm referencing and criterion referencing to make recommendations about thresholds for good fidelity to the wraparound principles as assessed by the WFI-3. To do so, we employed a two-pronged approach: (1) Examined the national WFI-3 dataset, to determine norms for a sample of wraparound programs; and (2) Reviewed studies that collected fidelity data as well as data for one or more external criteria variables proposed to be associated with wraparound fidelity. By considering the results of these two exercises together, we hoped to be able to shed light on what "good enough" wraparound fidelity scores may be.

Method

Measure

The WFI-3, is structured interview that assesses adherence to 11 core principles of wraparound (e.g., Family Voice and Choice, Individualized, Natural Supports, Team-Driven). Four items serve as indicators for each element, with responses ranging from 0 (*low fidelity*) to 2 (*high fidelity*). The resource facilitator and caregiver forms each contain 44 items while the youth form includes 32 items (the youth form of the WFI includes only 8 of the 11 elements). Total scores for each of the respondents are converted to a percent of total possible score (88 for the resource facilitator and caregiver, 64 for the youth). An overall fidelity score is also calculated that combines reports of the three respondents. Validation studies of the WFI have found adequate test-retest reliability and internal consistency for WFI total scores, as well as evidence for convergent and criterion-related validity.

Procedure

Norm-referencing exercise. WFI-3 data for 10 programs or communities in nine states nationally who used the WFI-3 were analyzed (total N=667 families; See Table 1). Means and standard deviations were calculated and plotted for these 10 programs. In addition, analyses of variance with post-hoc contrast effects were conducted to determine variability across program sites and what scores represented significant differences.

Insert Table 1 about here

Criterion-referencing exercise. To assess how WFI-3 scores relate to external criteria, a review was conducted of studies published, in press, or presented at a major conference that presented WFI-3 data for two or more groups that differed with respect to an external criterion. Five studies were included. Two were evaluation studies of wraparound vs. non-wraparound control groups that included WFI-3 scores for wraparound vs. non-wraparound comparison or control groups (Peterson et al., 2004; Ferguson, 2004). One study presented WFI-3 data for a sample of programs with poorer vs. better system and organizational supports for wraparound (Bruns, Suter, & Leverentz-Brady, 2004). Another study presented WFI-3 data for wraparound facilitators with poorer vs. better child outcomes achieved (Rast et al, 2004). The final study presented WFI-3 data for three stages of a program, whereby wraparound implementation support increased at each stage (no training, training, training + coaching) (Rast & VanDenBerg, 2004).

Results

Norm-referencing

Results of the norm-referencing exercise found that mean Overall WFI-3 scores for the 10 sites ranged from 72.2% to 80.1%, with a mean for all families assessed of 76.7% (SD = 2.3) (See Table 2). Results of ANOVAs demonstrated that sites scoring under 74% overall fidelity were significantly different from sites scoring over 79%. Individual respondents' mean total fidelity scores were 80.5% (SD = 3.2) for Resource Facilitators, 73.7% (SD = 4.2) for Caregivers and 73.6% (SD = 3.5) for Youths.

Insert Table 2 about here

Criterion-referencing

The two studies that assessed fidelity for both wraparound and non-wraparound comparison groups found significant between-group differences in WFI-3 scores, with WFI-3

overall fidelity scores at 60% and 64% for non-wraparound groups, compared to 75% and 76% for wraparound groups. The study of WFI-3 scores for programs or sites with poorer and greater supports for wraparound (as assessed by a standardized instrument) found that sites with poorer supports achieved mean WFI-3 scores of 72% versus 84% for sites with greater supports for wraparound. The study examining WFI-3 scores for individual wraparound facilitators found that facilitators whose clients achieved poorer outcomes achieved mean WFI-3 scores of 72%, compared to 87% for facilitators that achieved better outcomes. Finally, the longitudinal study of one community that received different implementation supports over time showed that mean WFI-3 scores increased from 64% pre-training, to 72% post-training, to 86% after implementation of both training and coaching.

Discussion

The goal of the current study was to examine WFI-3 scores from many different sources of data, in order to “bootstrap” guidelines for interpreting fidelity scores. To help interpret the results, we have presented the mean WFI-3 scores from the studies reviewed as well as the national dataset for the relevant conditions in Figure 1.

Insert Figure 1 about here

As shown, there is a discernable pattern whereby WFI-3 scores increase as greater supports for implementing the model are provided. Non-wraparound comparison conditions and a program not get formally trained demonstrated overall fidelity scores under 65%. Meanwhile, WFI-3 scores ranged from 72% to 76% for “wraparound as usual” conditions, such as wraparound programs with fewer system supports, wraparound facilitators whose children achieved poorer outcomes, wraparound with training only (not coaching), and wraparound groups from the evaluation studies. However, WFI-3 scores were found to be 84% to 87% for wraparound sites with better system supports, facilitators for whom children experienced better outcomes, and wraparound with both coaching and training.

WFI-3 scores across sites in our national WFI-3 database showed significant variability, but still fell logically within the pattern, ranging from 72% to 80% overall, with a mean of 76.7%.

By combining data from these norm-referenced and criteria-referenced approaches, we can begin to make some educated guesses about what represents adequate and good fidelity. As shown in Figure 1, we have proposed fidelity thresholds at 65%, 75%, and 85% overall fidelity scores. Clearly, scores below 65% are unlikely to represent true wraparound, as non-wraparound comparison groups and pre-training wraparound programs score in this range. Wraparound programs with poorer supports or that achieve poorer outcomes tend to fall between 65% and 75% fidelity, and thus are referred to as achieving “borderline” wraparound fidelity. The mean overall fidelity for the national dataset was found to be 77%, and half of the national sites scored above 75% (as did both wraparound programs in the evaluation studies), so we propose that this threshold represents “adequate” fidelity. Finally, we have proposed 85% and above as “high fidelity” because conditions of better supports or better outcomes were found to score at or above this threshold.

The benchmarking exercise we have undertaken is somewhat subjective and has several limitations. For example, different sites and studies used different methods for collecting WFI-3

data, which may have influenced scores. We also do not know much about wraparound implementation in most of the sites that were included. Nonetheless, we expect that the results of this analysis and the thresholds that have been set will be useful for programs as well as for researchers who use the WFI-3. Programs, communities, or researchers seeking to interpret their WFI-3 scores may also find the mean scores for individual respondents useful. Ultimately, the best information about what communities should be striving to achieve with respect to fidelity will come from more research on the relationship between administering the wraparound process and child and family outcomes.

References

- Bruns, E.J., Burchard, J.D., Suter, J.C., & Force, M.D. (2005). Measuring fidelity within community treatments for children and families. In Epstein, M., Duchnowski, A., & Kutash, K. (Eds.) *Outcomes for Children and Youth with Emotional and Behavioral Disorders and their Families*, vol. 2, 175-197. Austin, TX: Pro-ED.
- Bruns, E.J., Suter, J.C., Burchard, J.D., Force, M., & Leverentz-Brady, K. (2004). Assessing fidelity to a community-based treatment for youth: the Wraparound Fidelity Index. *Journal of Emotional and Behavioral Disorders*, 12, 69-79.
- Bruns, E.J., Suter, J.S., Force, M.D., & Burchard, J.D. (2005). Fidelity to the wraparound process and its association with child and family outcomes. *Journal of Child and Family Studies*, 14, 521-534.
- Bruns, E.J., Suter, J.D., & Leverentz-Brady, K. (2004). A national portrait of wraparound implementation. In C. Newman, C. Liberton, K. Kutash, & R.M. Friedman (Eds.), *The 16th Annual Research Conference Proceedings: A System of Care for Children's Mental Health*. Tampa: University of South Florida, Florida Mental Health Institute Research and training Center for Children's Mental Health.
- Bruns, E.J., Walker, J.S., VanDenBerg, J.D., Rast, J., Osher, T.W., Miles, P., Adams, J., & National Wraparound Initiative Advisory Group (2004). *Phases and activities of the wraparound process*. Portland, OR: National Wraparound Initiative, Research and Training Center on Family Support and Children's Mental Health, Portland State University.
- Burns, B.J., Hoagwood, K., & Maultsby, L.T. (1998). Improving outcomes for children and adolescents with serious emotional and behavioral disorders: Current and future directions. In Epstein, M., K. Kutash, and A. Duchnowski (Eds.), *Outcomes for Children and Youth with Behavioral and Emotional Disorders and their Families*. Austin, TX: Pro-Ed.
- Ferguson, C. (2004). California's title IV-E child welfare waiver demonstration project evaluation: An analysis of wraparound in Alameda County. In C. Newman, C. Liberton, K. Kutash, & R.M. Friedman (Eds.), *The 17th Annual Research Conference Proceedings: A System of Care for Children's Mental Health*. Tampa: University of South Florida, Florida Mental Health Institute Research and training Center for Children's Mental Health.
- Peterson, C.R., Gruner, L., Earnest, L., Rast, J., & Abi-Karam, N. (2004). Comparing functional outcomes of wraparound and traditional mental health and child welfare services. In C. Newman, C.J. Liberton, K. Kutash, & R.M. Freidman (Eds.), *The 16th Annual Research Conference Proceedings, A System of Care for Children's Mental Health: Expanding the Research Base*. Tampa: University of South Florida, The Louis de la Parte Florida Mental Health Institute, Research and Training Center for Children's Mental Health, 307-311.
- Rast, J., Peterson, C.R., Earnest, L., & Mears, S. (2004). Service process as a determinant of treatment effect – the importance of fidelity. In C. Newman, C.J. Liberton, K. Kutash, & R.M. Freidman (Eds.), *The 16th Annual Research Conference Proceedings, A System of Care for Children's Mental Health: Expanding the Research Base*. Tampa: University of

- South Florida, The Louis de la Parte Florida Mental Health Institute, Research and Training Center for Children's Mental Health, 311-315.
- Rast, J. & VanDenBerg, J. (2004). Certification of facilitators as a method for increasing wraparound fidelity. . In C. Newman, C.J. Liberton, K. Kutash, & R.M. Freidman (Eds.), *The 16th Annual Research Conference Proceedings, A System of Care for Children's Mental Health: Expanding the Research Base*. Tampa: University of South Florida, The Louis de la Parte Florida Mental Health Institute, Research and Training Center for Children's Mental Health, 315-320.
- Salyers, M.P., Bond, G.R., Teague, G.B., Cox, J.F., Smith, M.E., Hicks, M.L., Koop, J.I. (2003). Is it ACT yet? Real world examples of evaluating the degree of implementation for assertive community treatment. *Journal of Behavioral Health Services*, 30(3), 304-320.
- U.S. Public Health Service. (1999). *Mental health: A report of the Surgeon General*. Rockville, MD: U.S. Department of Health and Human Services Administration, Center for Mental Health Services, National Institutes of Health, National Institute of Mental Health.
- U.S. Public Health Service. (2001). *Youth violence: A report of the Surgeon General*. Rockville, MD: U.S. Department of Health and Human Services Administration, Center for Mental Health Services, National Institutes of Health, National Institute of Mental Health.
- Walker, J.S., Bruns, E.J., Adams, J., Miles, P., Osher, T.W., Rast, J., VanDenBerg, J.D. & National Wraparound Initiative Advisory Group (2004). *Ten principles of the wraparound process*. Portland, OR: National Wraparound Initiative, Research and Training Center on Family Support and Children's Mental Health, Portland State University.
- Walker, J. S., Koroloff, N., & Schutte, K. (2003). *Implementing high-quality collaborative individualized service/support planning: Necessary conditions*. Portland OR: Research and Training Center on Family Support and Children's Mental Health.

Table 1. Number of families and WFI respondents from each community in the national WFI-3 sample

Site	N Families	N Facilitators	N Caregivers	N Youths
Indiana	49	44	39	23
Nebraska	366	339	226	212
Pennsylvania	18	17	14	12
Massachusetts	74	74	71	29
Minnesota 1	26	26	23	12
North Carolina	22	22	22	16
Missouri	34	34	32	27
Nevada	31	31	31	23
Florida	27	24	17	5
Minnesota 2	21	19	15	8
Total Ns	667	622	490	367

Table 2. Overall WFI-3 fidelity scores for the 10 national study sites.

Site	Site mean WFI-3 score	SD	Between-site differences
1	0.722	0.11	a
2	0.735	0.09	a
3	0.735	0.07	a
4	0.751	0.09	ab
5	0.753	0.12	ab
6	0.794	0.09	ab
7	0.795	0.10	ab
8	0.797	0.11	b
9	0.800	0.07	b
10	0.801	0.09	b
Mean for all sites	0.767		

Note. Sites have been de-identified and presented in rank order. Between site differences as assessed via post-hoc contrast effects are indicated by coefficients with different letters. ANOVA result: $F(9,656) = 5.95, p < .0001$.

Figure 1. Summary of results of criterion-related and norm-related exercises, with proposed WFI-3 fidelity standards.

