HHS Public Access

Author manuscript

Adm Policy Ment Health. Author manuscript; available in PMC 2023 July 05.

Published in final edited form as:

Adm Policy Ment Health. 2020 November ; 47(6): 932–934. doi:10.1007/s10488-020-01069-4.

Commentary: Challenges and Opportunites in the Assessment of Fidelity and Related Constructs

Shannon Wiltsey Stirman, PhD

National Center for PTSD and Stanford University, 795 Willow Rd, NC-PTSD Menlo Park, CA 94025

Introduction

In this special issue, Bond and Drake (2019) capture some of the key challenges and considerations in evaluation of these measures. It is laudable that this issue focuses on a careful evaluation of the psychometric properties of fidelity measures, as many have not been closely evaluated. Fidelity is considered an implementation outcome, but it may also influence clinical outcomes. As such, fidelity measures have been developed and used in intervention process and outcome research, but the psychometric properties of many have not been examined closely. Reliable, valid measurement is critical to establishing what levels of fidelity are needed both to consider a program fully implemented, to understand factors that are necessary and sufficient for desired outcomes (Ruud, Hoifodt et al., 2020). While it is relatively easy to describe high, moderate, and low fidelity based on presence or absence of specific elements or on the quality of what was provided, it is more challenging to establish what specific elements, doses of exposure, and level of fidelity are sufficient to produce the outcomes we seek. It may also be important to distinguish between programs and interventions, while recognizing that nested within each are numerous specific factors that must be assessed. Many components are nested within a broader program or intervention, and often the elements that are most essential have not been empirically established. Determination of associations between intervention outcomes and intervention fidelity (to specific components or the entire intervention) careful measurement at multiple timepoints (Webb et al., 2010), under circumstances in routine care settings that make it challenging to isolate or experimentally manipulate specific elements of the program.

Levels of Fidelity

Fidelity to Program Characteristics. Programs comprise multiple components, and require fidelity measures to assess the presence of each of these elements. Such measures tend to assess the degree to which each component is in place, and the concept of adherence or

Terms of use and reuse: academic research for non-commercial purposes, see here for full terms. https://www.springer.com/aam-terms-v1

650-493-5000 x20007, Sws1@stanford.edu.

Publisher's Disclaimer: This Author Accepted Manuscript is a PDF file of an unedited peer-reviewed manuscript that has been accepted for publication but has not been copyedited or corrected. The official version of record that is published in the journal is kept up to date and so may therefore differ from this version.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

extensiveness may be more central for programs than quality or competence. For example, presence or absence of key staff members, policies, or activities indicates whether these aspects of the program have been implemented as intended. In some programs, one or more specific intervention (e.g., Dialectical Behavior Therapy skills group, Seeking Safety) is a nested within the program. For example, in Illness Management and Recovery (Egeland et al., 2019), cognitive behavioral techniques, coping skills, and relapse prevention training are all required elements. These interventions, in turn, may consist of numerous elements that must be delivered competently to be considered to be fully or appropriately implemented.

Intervention fidelity. Fidelity assessments for interventions such as cognitive behavioral strategies are typically developed in clinical trials, and most commonly include assessments of both adherence (whether or not a component was delivered) and competence (the degree of skill with which it was provided). When the entire intervention, or elements of it, are nested within a program (e.g., Egeland et al., 2019), then, it is important to assess whether these interventions, as key components of the broader program, are present. However, there are several challenges to the feasibility and scalability of these components. Fine-grained assessments of intervention fidelity at this level can be challenging, as they often require time-consuming observation or reliable self-report, which can be elusive (which cognitive behavioral strategies or coping skills were emphasized? Were they taught, used and reinforced skillfully and appropriately?). They may take place in the context of scheduled group interventions or classes, but they may also be woven through and reinforced throughout the day (Riggs & Creed, 2017), which can make observer assessment difficult. However, some data suggest that observer ratings are more reliable than supervisor or provider self-report data (Caron et al., 2019); although other studies have found that providers may be able to accurately report on the less nuanced aspects of fidelity (Ward et al., 2013). Additionally, numerous observer ratings—more than are feasible on a large scale-- are required to ensure a stable estimate of the interventionist's level of fidelity (Dennhag et al., 2012). As noted by Bond et al (2019), rater calibration may be particularly challenging for such items, as raters need to understand what competent delivery of the components looks like. Initial and ongoing calibration can be fairly labor intensive, but necessary to ensure consistent standards and accurate feedback. Such an investment may be important when high-stakes decisions are made—such as certifications of programs or therapists, funding, and policy decisions. Additionally, when there is a clearly established link between fidelity to the intervention and clinical outcomes, it may be particularly important to monitor and support fidelity. In fact, there is some evidence that observation and fidelity monitoring may improve clinical and implementation outcomes (Robbins et al., 2019; Aarons et al., 2010).

Supplementing Fidelity Assessment with Other Measures

Implementation and Quality Measures. Often the measures that are developed for clinical trials are used for assessment once interventions are implemented in routine practice. However, these scales often neglect additional factors that are baked into intervention research, which may themselves influence the degree of fidelity, or the clinical outcomes. Ruud, Hoifoldt, and colleagues' (2020) finding that organizations are more likely to establish policies related to implementation than they are to fully implement new programs

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

suggests the need for support and structure around the implementation itself. Heiervang and colleagues (2020) point out that fidelity assessment of specific practices does not include measurement of individualization and quality improvement that might influence program outcomes. These activities, which often accompany the intervention or program itself in clinical trials, are important for ensuring quality, consistency, and appropriate care (Lyon, Stanick, & Pullman, 2018). Considering their influence on, and interaction with fidelity, may advance the field's understanding of how the process vs. content of implementation impacts program or intervention outcomes. In fact, programs may have better outcomes when we begin to consider these elements as essential as the elements of the specific program or intervention.

Adaptation. At both a program level, if applicable, and at the level of a specific psychological intervention, numerous factors may impact capacity and ability to provide the program as originally intended. Some circumstances will require adaptation. Adaptation can take many forms, ranging from changes in setting or format to the number or type of personnel who deliver an intervention. Changes to the content of interventions can range from minor tailoring to changing timing, or adding, removing, or substituting elements. Adaptations can be consistent or inconsistent with fidelity (Stirman et al., 2015; Marques et al., 2019). Some adaptations appear to enhance outcomes (Stirman et al., 2017; Marques et al., 2019). Others, particularly removal of key elements, are inconsistent with fidelity and may lead to decreases in the effectiveness of the program or interventions. Key to determining whether an adaptation is fidelity-consistent is whether core elements of the component are changed. However, more recently, implementation scientists have begun to look beyond the form of an element of the intervention or program, to its actual function or goal (Jolles, Legnick-Hall, & Mittman, 2019). If the function is preserved, program or intervention components can take many forms, as long as the key function has been preserved. For example, if the function or goal of psychoeducation in Illness Management and Recovery (Ruud, Hoifodt, et al., 2020) is to ensure that the consumer understands their condition and how to manage it, psychoeducation could in theory take many forms (peer-led groups, a provider-led orientation meeting, a game, or videos that are watched and then discussed) and could be adapted to accommodate local constraints and consumer preferences, as long as the goal is met. Supplementing fidelity assessment with a measure of adaptations that occur when provided in routine care settings, and examining it in conjunction with evaluation data provides opportunity for learning about what the core functions or elements of interventions actually are essential within different contexts, and which forms are feasible and effective (Stirman, Miller, & Baumann, 2019; Miller, Stirman, Baumann, 2020). As a result of such evaluation, fidelity measures (either decision rules for each item, or the items themselves) may require updating to reflect any new knowledge. This process will ensure that fidelity measures that were developed for the purposes of research reflect the realities and context of routine care.

Conclusion

This special issue presents exemplars of the type of rigorous evaluation that has been lacking for many fidelity measures. Collectively, the articles demonstrate the many considerations that must be made to understand whether key components interventions as they are

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

implemented in communities. Ongoing program evaluation, refinement of these measures, and assessment of complementary constructs will allow the field to continue to advance our understanding of the role of fidelity in successful implementation.

References

Bond GR, Drake RE Assessing the Fidelity of Evidence-Based Practices: History and Current Status of a Standardized Measurement Methodology. Adm Policy Ment Health (2019). 10.1007/s10488-019-00991-6

Caron EB, Muggeo MA, Souer HR, Pella JE, & Ginsburg GS (2019). Concordance between clinician, supervisor and observer ratings of therapeutic competence in CBT and treatment as usual: does clinician competence or supervisor session observation improve agreement?. fs, 1–14. DOI: 10.1017/S1352465819000699 [PubMed: 31998581]

Dennhag I, Connolly Gibbons MB Barber JP, Gallop R, & Crits-Christoph P (2012) How many treatment sessions and patients are needed to create a stable score of adherence and competence in the treatment of cocaine dependence?, Psychotherapy Research, 22:4, 475–488, DOI: 10.1080/10503307.2012.674790 [PubMed: 22449079]

Egeland KM, Heiervang KS, Landers M et al. Psychometric Properties of a Fidelity Scale for Illness Management and Recovery. Adm Policy Ment Health (2019). 10.1007/s10488-019-00992-5

Guérin E, Dupuis JP, Jacob JD, & Prud'homme D (2019). Incorporating a Physical Activity Program into an Assertive Community Treatment Team: Impact and Strategies. Community mental health journal, 55(8), 1293–1297. 10.1007/s10597-019-00440-6 [PubMed: 31290032]

Jolles MP, Lengnick-Hall R, & Mittman BS (2019). Core functions and forms of complex health interventions: A patient-centered medical home illustration. Journal of general internal medicine, 34(6), 1032–1038. DOI: 10.1007/s11606-018-4818-7 [PubMed: 30623387]

Kühne F, Meister R, Maaß U et al. How Reliable Are Therapeutic Competence Ratings? Results of a Systematic Review and Meta-Analysis. Cogn Ther Res 44, 241–257 (2020). 10.1007/s10608-019-10056-5

Lyon AR, Stanick C, & Pullmann MD (2018). Toward high-fidelity treatment as usual: Evidence-based intervention structures to improve usual care psychotherapy. Clinical Psychology: Science and Practice, 25(4), e12265.

Marques L, Valentine SE, Kaysen D, Mackintosh MA, De Silva D, Louise E, … & Wiltsey-Stirman S (2019). Provider fidelity and modifications to cognitive processing therapy in a diverse community health clinic: Associations with clinical change. Journal of Consulting and Clinical Psychology, 87(4), 357. 10.1037/ccp0000384 [PubMed: 30883163]

Miller CJ, Wiltsey-Stirman S, & Baumann AA (2020). Iterative Decision-making for Evaluation of Adaptations (IDEA): A decision tree for balancing adaptation, fidelity, and intervention impact. Journal of Community Psychology 10.1002/jcop.22279

Riggs SE, & Creed TA (2017). A model to transform psychosis milieu treatment using CBT-informed interventions. Cognitive and Behavioral Practice, 24(3), 353–362. 10.1016/j.cbpra.2016.08.001

Robbins MS, Waldron HB, Turner CW, Brody J, Hops H, & Ozechowski T (2019). Evaluating supervision models in functional family therapy: Does adding observation enhance outcomes? Family Process, 58(4), 873–890. 10.1111/famp.12399 [PubMed: 30339285]

Ruud T, Drivenes K, Drake RE et al. The Antipsychotic Medication Management Fidelity Scale: Psychometric properties. Adm Policy Ment Health (2020). 10.1007/s10488-020-01018-1

Ruud T, Høifødt TS, Hendrick DC et al. The Physical Health Care Fidelity Scale: Psychometric Properties. Adm Policy Ment Health (2020). 10.1007/s10488-020-01019-0

Stirman SW, Baumann AA, & Miller CJ (2019). The FRAME: an expanded framework for reporting adaptations and modifications to evidence-based interventions. Implementation Science, 14(1), 58. 10.1186/s13012-019-0898-y [PubMed: 31171014]

Ward AM, Regan J, Chorpita BF, Starace N, Rodriguez A, Okamura K, … & Research Network on Youth Mental Health. (2013). Tracking evidence-based practice with youth: Validity of the

MATCH and Standard Manual Consultation Records. Journal of Clinical Child & Adolescent Psychology, 42(1), 44–55. 10.1080/15374416.2012.700505 [PubMed: 22809135]

Webb CA, DeRubeis RJ, & Barber JP (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. Journal of Consulting and Clinical Psychology, 78(2), 200–211. 10.1037/a0018912 [PubMed: 20350031]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript