

DEFINING TREATMENT INTEGRITY: ADHERENCE AND COMPETENCE IN PRACTICE

Chair: Elizabeth Gifford, PhD, Center for Health Care Evaluation, VA Palo Alto Health Care System and Stanford University School of Medicine

Case Presenters: John Baer, PhD, University of Washington, Nancy Roget, MA, University of Nevada, Reno, and Craig Rosen, PhD, VA Palo Alto Health Care System and Stanford University School of Medicine

Discussant: Ken Weingardt, PhD, Center for Health Care Evaluation, VA Palo Alto Health Care System and Stanford University School of Medicine

Think Tank Overview

Assessing program fidelity, or the degree to which an intervention in the field follows a protocol developed and evaluated in clinical trials, is an important part of the effort to disseminate research-based protocols. Treatment integrity or fidelity was originally a check on the independent variable used to identify whether treatments in randomized controlled trials were delivered as intended. Fidelity in this sense has three components: (a) adherence, defined as the extensiveness of dosage of model-prescribed intervention techniques implemented in session and the absence of proscribed practices; (b) treatment differentiation, or the uniqueness of the treatment relative to comparison treatments provided in the trial; and (c) competence, or the quality or skillfulness with which the therapist delivered the treatment.

The mental health treatment field has begun to shift its focus from disseminating Empirically Supported Treatments (ESTs), which establish the efficacy of specific treatments in randomized controlled trials, to implementing Empirically Based Practices (EBPs), in which evidence-based interventions are delivered within a context that includes the provider's clinical expertise and patient preferences. The essential goals of EST fidelity in treatment development are not identical to the goals of EBP fidelity. For example, practice settings require clinical decision making that is excluded from the more controlled world of efficacy trials, where patients are pre-screened in order to achieve diagnostic purity for research purposes, ongoing assessments are conducted by a research team but rarely used to inform treatment, supervision is provided as part of a research protocol, etc. Comparing fidelity evaluation in efficacy versus practice settings indicates that their purposes both overlap (e.g., verifying delivery of the independent variable in effectiveness or efficacy trials, tracking/reducing proscribed practices), and differ (e.g., treatment differentiation is less important in the practice setting, clinical decision making is critical, and patient progress is paramount).

This Think Tank session explored whether traditional definitions of fidelity or integrity are adequate for use in treatment practice settings and appropriate for the purposes of implementation science. A panel of EBP implementation experts working in both practice and research settings presented three brief case histories: motivational interviewing (MI) treatment fidelity research, technology transfer of substance use disorder (SUD) treatments, and a national post traumatic stress disorder (PTSD) treatment training initiative within the Veterans Administration (VA). Future directions in implementation fidelity include (a) providing guidelines for adaptation of EBPs during implementation, for example, identifying active ingredients in practice (i.e., EBP components linked to patient outcomes) that should be included in adaptations and accompanying fidelity evaluations, (b) studying the impact of program, supervision, and therapist fidelity on patient outcomes in the field, and (c) logistical questions about how to feasibly conduct process-based assessment, e.g., systematically exploring methods for gathering work sample data from clinicians, validating technological tools such as web based clinician assessment, etc. Serious efforts to improve fidelity must address existing resource constraints in funding and infrastructure that limit programs' ability to implement EBPs and conduct fidelity evaluations.

Adherence and Adaptation

Thoughtful adaptation of evidence-based treatment protocols so that they are consistent with the provider's clinical judgment, the patient's individual preferences, and the constraints of the practice setting, appears to be critical to successful implementation. According to two speakers, clinicians are likely to make their own changes over time, and the degree to which this is a problem remains a question. Several audience members and panel members discussed barriers to implementing programs and fidelity evaluations without adaptation. For example, participants noted that clients who are typically served in practice are not only different than those who typically participate in research studies, but may change rapidly in response to organizational and societal factors. Another participant noted that in an effort to disseminate MI statewide in a primary care chronic illness treatment system, the intervention was just too complex to be implemented successfully, and that interventions need to be simpler and more pragmatic if we hope to promote adoption.

It would appear that the critical question is not whether to adapt treatment protocols for delivery in clinical practice, but how to adapt them without compromising the validity of the intervention. How much can a protocol be changed and remain effective? At what point is it too removed from the tested protocol to be an empirically based approach? Speaking colloquially, "How far can a protocol bend before it breaks?" Answering this question requires that we know which treatment components are non-negotiable active ingredients and which are relatively inert.

Participants described two research strategies that may help to identify the active ingredients of evidence-based treatments. This first is a class of studies evaluating the minimum level of fidelity necessary to achieve desired outcomes (i.e., "How low you can go?"). As one speaker noted, with clinicians making changes in practice we may want to look for "U-shaped fidelity curves" in order to identify the degree to which providers can make adaptations and still have good outcomes. A second approach consists of dismantling studies focusing on fidelity to necessary active ingredients or core components, including core process related interaction styles such as empathy (e.g., "MI style" as the critical component of motivational interviewing treatment).

Measuring Fidelity, Skill and Competence

Skill is defined "as the extent to which the therapists conducting the interventions took the relevant aspects of the therapeutic context into account and responded to these contextual variables appropriately" (Waltz, Addis, Koerner & Jacobson, 1993, p.620). A critical aspect of fidelity assessment is the skillful application of research based protocols in the context of the preferences, needs and progress of specific patients. The importance of skillful response to the patient was in the earliest definitions of competence (e.g., Waltz et al., 1993), and is emphasized in recent studies indicating that even adherence must be done flexibly and skillfully in order to maximize patient outcomes. The practice setting is a rich and changing context requiring ongoing skillful response and adaptation on the part of the clinician, and thus appears to rely even more on competence than adherence.

Because competence involves a skillful response to a particular patient in context, it is intrinsically process-oriented. Therefore, measuring competence at the point of service requires observing work samples. William Miller, the developer of the MI approach to the treatment of addictive disorders, proposed a system of measuring clinical proficiency based on the review of session tapes, with several process-based criteria for therapists including: global ratings of empathy and congruence with the spirit of MI, talking less than the client, using reflections twice as often as questions, using reflections that are more often complex than simple, and asking open ended questions more often than closed questions.

Systems for coding MI fidelity from session tapes include the Motivational Interviewing Skills Code (MISC; Miller, 2000), the Motivational Interviewing Treatment Integrity scale (MITI; Moyers et al., 2005). The Behavior Change Counseling Index, (BECCI; Lane et al., 2005), and the Independent Tape Rater Scale (ITRS, a modification of the Yale Adherence and Competence Scale (YACS); Ball et al., 2002). Alternative methods of assessment of MI competence include: Questionnaire (HRQ), Video Exams (VASE-R), Computer Exams (CASPI), Standardized Patient Interviews (SPIs). Video and computer-based programs for assessing skills are in development. For example, a video exam where clinicians view video vignettes and then write down what they would say next, or a computer-based system (the CASPI) where the therapist actually speaks aloud into the computer what they would say next. However, the validity and generalizability of these technology-based skill assessment methods have yet to be established.

The most sophisticated, valid and reliable coding system is useless unless clinicians are willing to provide audiotapes of their counseling sessions. The MI Assessment Supervision Tools to Enhance Proficiency (MIA-STEP), which comes out of the Clinical Trials Network, begins by asking the pragmatic question of how to gather and analyze tapes of sessions so that they can be rated for competence. Unfortunately, the answers to that question are generally unsatisfactory or incomplete. It is possible, albeit strenuous, to reliably train raters to code and rate tapes with approximately 40 hours of training. However, work samples are difficult and costly to obtain. The field simply does not have systems in place to collect these kinds of data. This is a significant barrier to evaluating fidelity of interventions.

The Think Tank discussion focused on the challenges of acquiring tapes from practicing clinicians in the field. One participant affiliated with the California Institute for Mental Health reported successfully obtaining tapes from clinicians by focusing on building trust with the clinicians. This was accomplished by providing a lot of positive feedback and working through barriers such as fears about “gotcha” mentality, concerns related to job security, etc. Others note a larger culture change is required, where clinicians see providing tapes as a basic part of their job description. In order for this to happen, however, it is necessary to have supportive management and system level involvement.

A related discussion concerned potential threats to the validity of the audiotaped work samples. Some participants suggested that in order to facilitate clinician compliance, clinicians should be permitted to choose their own samples even though they may select samples that are not entirely representative of their weaknesses. Even if therapists choose samples of their best work, these participants argued, it is fairly safe to assume that high level skills are likely to generalize between patients or populations. Other participants suggested that allowing clinicians to pick their best tapes may prevent supervisors from detecting whether they are also doing things that could be iatrogenic, which in adherence or fidelity terms are described as proscribed practices. This may turn out to be more important than we realize. Indeed, ultimately the field may not be as interested in adherence in terms of treatment differentiation, but adherence in terms of avoiding proscribed practices that lead to bad outcomes. Lastly, it was noted that just as researchers hope to sample representative patients, it may also be possible to sample representative clinicians.

Focusing on competence raises some potentially sensitive issues, including the fact that not everyone can perform interventions well. This is particularly problematic, as one speaker noted, because more elaborate training programs do not always lead to better outcomes, particularly in treatments or programs that rely upon intervention skills such as empathy and interpersonal concordance. This led to a discussion of the simple question “what makes a good intervener?” A researcher from Johns Hopkins implementing a community based parenting prevention program observed that while they tape their interveners and get very good adherence, they have observed that this is not sufficient for quality intervention. Competence or skill is essential.

Moving to a systems perspective, participant case studies described the challenges of measuring fidelity in large health care systems. Presenters agreed that the processes by which clinicians implement evidence-based practices will remain a black box until there are large scale organizational mechanisms to assess therapists' competence or skill. Several different organizational pathways to gathering fidelity data in the context of a large health care system were presented. Data can be gathered via supervision, either through clinician report to supervisors, or to recordings observed by supervisors. Clinician surveys on frequency of use, or chart reviews that would identify occasions of use, can be conducted (although these have validity issues, one speaker noted that asking clinicians whether they use a treatment is an important first step and a significant organizational accomplishment). Other options include sampling providers for quality assurance. This would be less laborious than trying to get feedback on every clinician or every patient, but how to do it remains a question, e.g., whether to require tapes, use a web-based system with standardized patients, etc. Another option is to require periodic recertification after training in which therapists must bring in tapes and get supervision (e.g., every two years). Lastly, in order to answer the question of competence and adaptation, one wants to be able to assess the impact of the treatment at the level of the individual patient, so tracking patient progress or proximal outcomes is optimal. If patients are not making progress as expected, this could trigger fidelity assessment.

The ultimate purpose of focusing on fidelity is to improve patient outcomes. Fidelity may occur at multiple levels, including fidelity to the treatment model, fidelity to a clinical supervision model, and fidelity to a program model. If fidelity to a program and supervision model do not lead to improved patient outcomes, then fidelity becomes irrelevant. Practicing clinicians are understandably more concerned with patient improvement than with following a treatment manual (adherence), or whether they are providing interventions from only one treatment at a time (treatment differentiation). The research field has also moved toward understanding proximal outcomes or mediators of treatment process, in part because understanding how to improve treatment requires understanding the processes or proximal outcomes through which treatments have their effects.

Do different levels of fidelity in practice settings lead to significant differences in patient outcomes? This requires careful assessment of the target of fidelity, e.g., fidelity to which components or principles of treatment? The target of fidelity should ideally be those elements of treatment that have been shown to improve patient outcomes. However, there has been little effort to link research on treatment mediators with research on clinician fidelity. This leads then to a second question: Does fidelity to different components or principles result in significant differences in patient outcomes? The critical research bearing on this issue is yet to be conducted. Integrating measures of patient progress on proximal outcomes into standard treatment may provide clinicians and treatment systems with feedback on adaptation, clinician competence, and successful implementation of EBPs.

Implementation realities

What's happening on the ground in our publicly funded treatment programs has a critical bearing on whether and how we can conduct fidelity evaluations at point of practice. Workforce, clinical oversight, and infrastructure needs are the most pressing barriers. The Bureau of Labor Statistics shows that of the 83,000 substance abuse counselors in the nation, the average hourly pay is \$17.28, for a total annual salary of \$35,950. Perhaps because these professionals work for low pay and minimal benefits, there are few young professionals entering the field. This problem is exacerbated by high turnover in the field, with 50.8% of practitioners reporting being in their current position 5 years or less (Mulvey, et al. 2003). Treatment agencies experience a 25% staff turnover rate per year (Gallon, et al., 2003). Teachers or nurses have about a 12% turnover rate, so the problem in SUD treatment is particularly serious.

Training workshops need to be followed by audiotape feedback and coaching in order to be effective. Even putting aside the pragmatic difficulties of obtaining audiotapes from counselors, costs for this more elaborated training can be quite prohibitive, e.g., one CBT workshop for 25 counselors with audiotape and follow up coaching is approximately \$30,000 (\$1,200/participant; McLellan et al., 2003). Furthermore, the SUD treatment infrastructure is troubled by a lack of computers, absence even of email addresses for counselors in many centers, and by a dearth of supervision. There is perhaps no other area of health care where clinicians work with so little supervision (Carroll & Rounsaville, 2007). If we want to support quality implementation and appropriate ongoing fidelity, we need to build up our foundation of clinical supervision, access to computers, and avenues for workforce training and development before we focus on implementing EBPs and evaluating implementation fidelity.

Researchers studying implementation in child mental health treatment have done important work developing system friendly approaches to disseminating EBPs, which provides some cause for optimism. These models involve some component of ongoing consultation or supervision, typically by telephone (other possibilities include supervision by web counseling on a centralized basis). However, it will require a fairly big culture change to get support for the right level of clinical supervision into the publicly funded treatment system. While the VA is a better resourced system, it still faces significant issues in providing quality ongoing supervision. These problems need to be addressed prior to a push to implement EBPs, if the hope is to promote quality interventions.

In general, implementation requires a more iterative, formative evaluation approach in order to overcome implicit research assumptions and clinician biases that serve as barriers. For example, the pharmaceutical treatment development model of efficacy to effectiveness and back again does not provide direct feedback loops via multiple iterations because of the costs of drug development. Psychosocial treatment implementation clearly needs more iterations in the field in order to identify barriers and resolve the system or management issues that are necessary conditions for fidelity evaluation.

As noted above, and described in work on diffusion of innovations (Rogers, 2003), reinvention is a common process when innovations are adopted. One of the conclusions of this Think Tank is that the field needs clinical practice trials that re-evaluate interventions in terms of critical ingredients or processes, including critical components of treatment and treatment delivery. Evaluating interventions formatively in the implementation context, as opposed to the development context, will provide an evidence base guiding adaptation or skilful application in practice, which is a necessary condition for evaluating implementation fidelity. Fortunately, positive developments including instrumental technological innovations and a growing awareness of the need for direct practice research may facilitate these efforts.

Acknowledgments

We would like to acknowledge the participants in this Think Tank discussion and express our appreciation for their remarks. Those interested in resources for treatment integrity methods in clinical trials are referred to Perepletchikova, Hilt, Chereji & Kazdin, 2009. Please address questions regarding this Think Tank to Elizabeth Gifford, Elizabeth.Gifford@va.gov.

Resources

Perepletchikova, F., Hilt, L.M., Chereji, E. & Kazdin, A.E. (2009). Barriers to implementing treatment integrity procedures: survey of treatment outcome researchers. *Journal of Consulting and Clinical Psychology, 77*(2), 212–218.