

Facing the Challenges of Implementation

Dean L. Fixsen¹ and Terje Ogden²

¹University of North Carolina at Chapel Hill, NC, USA,

²The Norwegian Center for Child Behavioral Development, Oslo, Norway

Ensuring that scientific findings bear fruit in terms of successfully employed real-world practical applications is not always an easy task. The challenges inherent in putting research knowledge into practice have contributed to the emergence of the field of implementation research, a young but quickly expanding discipline dealing with the “what,” “how,” and “who” of implementation: *What* is to be implemented, *how* is the task to be carried out, and *who* will do it? Implementation science has grown from an increasing awareness of the “science to service gap,” that is, concern about the often limited success of transferring research-based practices to ordinary service settings in fields as diverse as early childhood education, social services, mental health, employment services, or criminal justice. Awareness of this gap has inspired research efforts and numerous publications on aids and obstacles to successful implementation, and this topical issue of the *Zeitschrift für Psychologie* aims to contribute to the ongoing development of the field.

The empirical studies in this issue illustrate the complexities and challenges facing scientists conducting implementation research. Few standardized measurement instruments with defined psychometric qualities are available, and the contextual complexity of real-world practice makes it difficult to control for all potentially confounding variables. On the positive side, the articles illustrate the increasing sophistication of implementation research methodology, which allows for analyses of multivariate interactions and multilevel relationships. It is encouraging to see how long-term data collection can lead to more complex and quantitative analyses and meaningful measures (e.g., Brunk, Chapman, & Schoenwald, 2014).

Among the main challenges of implementation science is the disparity between optimal study designs needed to answer the research questions and the financial and other limitations present in the practice field. Many of the studies reported here had to accept less than ideal conditions by including relatively small samples (at the organizational level) and reliance on self-reports. Even if multi-informant

studies with large samples would be ideal, most have to settle for less. And even if sophisticated methods of statistical analyses may compensate for some of the shortcomings, most studies appear as pilot studies with preliminary findings. Adding to the challenges is the need to validate the measures of implementation components in order to produce robust and sustained outcomes.

Goense, Boendermaker, van Yperen, Stams, and van Laar (2014) is the first review of how measures of treatment integrity are used in outcome studies of evidence-based interventions for youths with externalizing behavior problems. The review is an illustration of how the field of implementation research is growing and how it is fast becoming an integral part of outcome studies. Increasing demands are put on the development and reporting of treatment integrity in clinical trials. Goense and colleagues found the assessment of treatment integrity in studies of evidence-based programs is rare. In particular, therapist competence was reported far less often than therapist adherence. The lack of competence measures makes it difficult to determine how differences in competence may have influenced differences in outcome effectiveness. In addition, many assessments of treatment integrity relied on self-reports, which precluded reports on training of raters and indicators of inter-rater reliability. The authors conclude that even if measuring treatment integrity is important, it is often missing in studies of interventions for youths with externalizing behavior problems. This weakens the confidence in statements about the association between treatment integrity and intervention outcomes. In future, researchers and journal editors will probably be more concerned about the measurement of treatment integrity as a moderator of clinical outcome studies, and more aware of the need to measure it in comparable ways across studies.

In their study, Brunk, Chapman, and Schoenwald (2014) take fidelity assessment to a new level. This study represents a major contribution to the next generation of implementation research and practice. The authors report the development of a measure of fidelity at the program

level to assess the quality of multisystemic therapy (MST) as used in typical provider organizations. Using a strong analytical approach, a single program score (PPI: The MST Program Performance Index) was calculated based on multiple indicators including treatment adherence, treatment completion, program operation guidelines, program capacity, clinical supervisor leadership, and stakeholder relationships. The predictive validity of the PPI score was supported by both client- and team-level outcomes. The authors conclude that the PPI may be a useful tool to index program level fidelity for comparative purposes and may be used as an additional tool for a decision support system. The PPI composite may be particularly useful in the process of large-scale implementations by measuring dimensions of fidelity beyond clinical adherence.

The study attests to the importance of multiple fidelity components to the quality outcomes for clients and to program sustainability. The description provides a roadmap for how other programs may approach the development of similar tools when they reach the appropriate stage of program development.

Klest (2014) examines the relationship between the number of PMTO (Parent Management Training Oregon model) practitioners working in an agency and implementation outcomes for those therapists. Using data from a survey study among a random selection of PMTO therapists in Norway, she finds that clustering PMTO therapists in the same agency may support successful program implementation. The most prominent differences in therapist-reported implementation outcomes were between those in agencies with three versus those in agencies with two practitioners. The findings suggest a critical mass of practitioners may make a substantial difference for sustaining the use of an evidence-based program in demanding service environments.

Romney, Israel, and Zlatevski (2014) provide a case example of implementation coupled with some data on costs and outcomes associated with organizational readiness. The authors examine the effects of agency-level implementation variation on the related outcomes of participant attrition and cost-effectiveness of an evidence-based parent training program (Triple P). In all, six agencies took part in the pilot study. Four agencies completed the exploration stage of implementation and met the project requirements. Prior to the training, administrators and staff from these four agencies participated in several site readiness activities to prepare them for the implementation process. Two other agencies participated as part of funding requirements and did not meet the project readiness requirements. In this preliminary study, the authors found the average cost-per-graduate was more than seven times higher for the two agencies that did not complete the readiness process compared to the four completing agencies. Moreover, the odds of Triple P participants graduating were 12 times greater for those in groups run by completing agencies. While this is a study with a small convenience sample, the measures used and the findings provide encouragement

for larger studies of the impact of creating readiness during early exploration work with agencies that are potential users of evidence-based programs.

Schultes, Stefanek, van de Schoot, Strohmeier, and Spiel (2014) investigate how the combined effect of two different aspects of program implementation – fidelity and participant responsiveness – influenced proximal outcomes of a school-wide program for the prevention of violence and bullying among students in Austria. Intervention fidelity measured the extent to which trainers implemented the intended program components, while participant responsiveness described the active application of the program contents by the target group. Multilevel analyses showed that teachers' self-efficacy was significantly enhanced in schools where the program was implemented with high fidelity. However, only teachers with high participation responsiveness (e.g., those who volunteered to train the students) significantly changed their behavior toward victims of bullying. Thus, distinguishing between these two aspects of implementation is necessary for a reliable interpretation of evaluation results. Different indicators of implementation may influence program outcomes differently and high implementation quality may be more relevant for some program outcomes than for others.

Rimehaug (2014) approaches the implementation process from quite another angle than most articles in this issue. We think that this is an angle that will resonate with many readers from the practice field as well as with researchers. Taking more of a practitioner perspective, the author describes how new practices were implemented and sustained over a 10-year period. Using a narrative approach, the author presents a case example that illustrates experiences, claims, and conclusions from the implementation literature. The aim of the study is to analyze how implementation strategies, setbacks, quality decay, challenges, and missing strategies as perceived by practitioners and purveyors are reflected in the implementation literature. Several generic recommendations relevant for practitioners, managers, and organizations are presented.

The articles in this issue target important topics of concern for researchers, policy makers, and practitioners alike. Central to several of the contributions is the need for improving the theoretical and operational definitions of constructs and the challenges of measuring them in reliable and valid ways. A recurring theme is therefore the need for more sophisticated ways of conceptualizing, measuring, and analyzing the implementation process.

Most of the studies deal with practical as well as methodological challenges to implementation science. They also deal with the questions of how study outcomes may be used in the process of planning new implementations, making practical decisions, and monitoring or evaluating outcomes at the program or individual level, including cost-effectiveness. As noted by these excellent articles, the field of implementation has gathered momentum and is prepared for rapid growth and development in the coming decades.

References

- Brunk, M. A., Chapman, J. E., & Schoenwald, S. K. (2014). Defining and evaluating fidelity at the program level: A preliminary investigation. *Zeitschrift für Psychologie*, 222, 22–29. doi: 10.1027/2151-2604/a000162
- Goense, P., Boendermaker, L., van Yperen, T., Stams, G.-J., & van Laar, J. (2014). Implementation of treatment integrity procedures: An analysis of outcome studies of youth interventions targeting externalising behavioral problems. *Zeitschrift für Psychologie*, 222, 12–21. doi: 10.1027/2151-2604/a000161
- Klest, S. K. (2014). Clustering practitioners within service organizations may improve implementation outcomes for evidence-based programs. *Zeitschrift für Psychologie*, 222, 30–36. doi: 10.1027/2151-2604/a000163
- Rimehaug, T. (2014). The ecology of sustainable implementation: A 10-year case history illustration. *Zeitschrift für Psychologie*, 222, 58–66. doi: 10.1027/2151-2604/a000166
- Romney, S., Israel, N., & Zlatevski, D. (2014). Effect of exploration-stage implementation variation on the cost-effectiveness of an evidence-based parenting program. *Zeitschrift für Psychologie*, 222, 37–48. doi: 10.1027/2151-2604/a000164
- Schultes, M.-T., Stefanek, E., van de Schoot, R., Strohmeier, D., & Spiel, C. (2014). Measuring implementation of a school-based violence program: Fidelity and teachers' responsiveness as predictors of proximal outcomes. *Zeitschrift für Psychologie*, 222, 49–57. doi: 10.1027/2151-2604/a000165

Dean L. Fixsen

National Implementation Research Network
FPG Child Development Institute
University of North Carolina at Chapel Hill
Campus Box 8040
Chapel Hill, NC 27517
USA
Tel. +1 919 966-3892
Fax +1 919 966-7463
E-mail dean.fixsen@unc.edu
