



The role of treatment acceptability, effectiveness, and understanding in treatment fidelity: Predicting implementation variation in a middle school science program



Joni M. Lakin*, David M. Shannon

Department of Educational Foundations, Leadership, and Technology, Auburn University, USA

ARTICLE INFO

Article history:

Received 19 January 2015
Received in revised form 25 May 2015
Accepted 9 June 2015
Available online 2 July 2015

Keywords:

Program evaluation
School-based evaluation
Evaluation methods
Fidelity of implementation

ABSTRACT

Evaluating a school-based program is a particular challenge when teachers implementing the program. Variations in implementation can be difficult to measure and predict. We developed measures to explore variation in treatment implementation and serve as predictors of variation in a school-based science education program. Based on previous work, we focused on assessing treatment acceptability, effectiveness, and understanding among teachers as critical determinants of variations in program implementation using self-reported and objective measures of implementation. All three constructs were associated with implementation of the program. Our measures of these constructs show promise for use in formative and summative evaluations. Our stratification of program elements with implementation predictors can provide a template for future exploration of treatment fidelity.

© 2015 Elsevier Ltd. All rights reserved.

Educators, educational researchers, and their funders are eager to understand why apparently effective educational interventions show so much variability in their success during scale up to widespread use. Researchers from a variety of fields have studied treatment fidelity and found that variations in how the program is implemented can often explain differences in treatment effectiveness across program implementations during scale up. Less is known about when and why variations occur, particularly for classroom interventions (Harn, Parisi, & Stoolmiller, 2013). This study focused on understanding implementation variation and treatment fidelity among teachers engaged in an instructional intervention in middle school science classrooms.

Treatment fidelity is the extent to which a program or treatment is enacted as originally intended and according to the design of program developers (Century, Rudnick, & Freeman, 2010; Gresham & Gansle, 1993; Mowbray, Holter, Teague, & Bybee, 2003; O'Donnell, 2008; Ruiz-Primo, 2006). In various literatures, it is also

referred to as fidelity of implementation (FOI), treatment integrity, dosage, or degree of program implementation (Ruiz-Primo, 2006). Studies of treatment fidelity address not only variations in how programs are adopted, but also how variations in use affect program effects on the target outcomes.

Treatment fidelity measures have been widely applied in public and mental health fields (Kazdin, 1980; Reimers, Wacker, & Koepl, 1987) and more recently in school psychology and assessment areas (particularly curriculum-based measurements, e.g., Allinder & Oats, 1997; O'Donnell, 2008). However, treatment fidelity is still relatively unexplored in terms of curriculum interventions that focus on content and pedagogy, particularly those outside of curriculum-based measurements and special education applications (O'Donnell, 2008). In some applications, treatment fidelity has been conceptualized as relatively straightforward dosage or degree of implementation. However, in other applications, especially educational interventions, treatment fidelity has been defined as a multidimensional construct including other issues such as how well the person implementing the program understands and can competently implement the program elements (Shulte, Easton, & Parker, 2009).

Classroom interventions are particularly challenging to evaluate and understand because teachers are responsible for carrying out critical components of the program while maintaining the teacher's own curriculum and professional standards. Multiple

Abbreviations: PCK, pedagogical content knowledge; POSITT, Pedagogy of Science Inquiry Teaching Test; STEM, science, technology, engineering, and mathematics; SECME, South Eastern Consortium for Minorities in Engineering.

* Corresponding author at: Department of Educational Foundations, Leadership, and Technology, Auburn University, 4036 Haley Center, AL 36849, USA. Tel.: +1 334 844 4930.

E-mail address: joni.lakin@auburn.edu (J.M. Lakin).

<http://dx.doi.org/10.1016/j.stueduc.2015.06.002>

0191-491X/© 2015 Elsevier Ltd. All rights reserved.

demands on a teacher's time means that they often cannot enact a program exactly as intended because of overlapping programs or competing classroom goals (Harn et al., 2013). Because of the complex nature of a school-based program, and the reduced control that program developers have over implementation, researchers and program developers need strategies to track the amount of implementation of the program exhibited by the teachers and to determine why teachers vary in their implementation (Swanson, Wanzek, Haring, Ciullo, & Mc-Culley, 2011).

Given the proliferation of science, technology, engineering, and mathematics (STEM) related school interventions, extending the literature on treatment fidelity to school-based STEM programs, as represented in this paper, will expand our knowledge not only of such interventions but also of evaluation theory. Our assessment plan drew from recent theory that emphasized the importance of using a variety of measures and a broad conceptualization of treatment fidelity in evaluating instructional interventions (Brandon, Young, Pottenger, & Taum, 2009; O'Donnell, 2008). We focused on assessing three constructs related to treatment fidelity—treatment acceptability, effectiveness, and understanding—among teachers as critical determinants of variations in program implementation. Our research questions focused on predicting variation in objective and self-reported implementation using measures of treatment acceptability, effectiveness, and understanding. We also explored how these four constructs (acceptability, effectiveness, understanding, and implementation) are best measured and used over time to provide formative (i.e., for program improvement) and summative (i.e., documenting program effects) feedback to program developers (Chambers, 1994).

1. Components and correlates of treatment fidelity

Treatment fidelity and predictive factors have been broken down into component parts in a number of ways by researchers seeking to explain its variability (e.g., Century et al., 2010; Mowbray et al., 2003; O'Donnell, 2008; Ruiz-Primo, 2006). The primary dimension of fidelity is implementation—the degree to which the program is actually enacted as intended by the program developers (Century et al., 2010). Theorists consider both structural and process qualities of implementation (Mowbray et al., 2003) dimensions that have been applied in practice (Century et al., 2010; Harn et al., 2013). Structural quality of implementation concerns more overt implementation—use of program activities/interventions—while process has to do with the quality of how the program activities/interventions are used or the quality of student–teacher interactions that arise from the program (O'Donnell, 2008). Both structural and process measures of implementation are ideal for understanding treatment fidelity in schools, although process measures are more difficult to assess and require careful definition of the program and implementation quality standards (Harn et al., 2013; Mowbray et al., 2003).

In delving into a multidimensional definition of implementation, Reimers et al. (1987) provided a helpful model for explaining why high-fidelity treatment implementation does not always occur, using factors that are external, but related, to implementation. Most relevant to educational interventions are the concepts of treatment acceptability, perceived effectiveness, and understanding, which Reimers et al. argued would moderate implementation and, therefore, mediate treatment effects. Treatment acceptability is based on the *perceived* appropriateness, fairness, reasonableness, and intrusiveness of a treatment to address a particular problem (Kazdin, 1980; Reimers et al., 1987). Effectiveness is the perception that the treatment will impact the problem or outcomes of interest. Treatment understanding refers to the program implementer's comprehension of general and specific components of the program or treatment and whether their comprehension is sufficient to

implement the treatment as intended by the program developers (Reimers et al., 1987). Notably, acceptability is influenced itself by the perception that there is a problem in need of a treatment, with more extreme problems being required to justify more extreme or disruptive treatments (Sterling-Turner & Watson, 2002). For school interventions, administration and community support as well as material and time costs are also determinants of acceptability (Broughton & Hester, 1993; Kurita & Zarbatany, 1991; Witt & Elliott, 1985).

In the classroom, teachers play a critical role in whether curricular programs have significant impacts because they determine whether and how much the program elements occur in the classroom (Allinder, 1996). In this context, acceptability and understanding act as gatekeepers for a program to have any impact on the target outcomes. Because teachers have so many demands on their time, they may not fully implement a program that they do not perceive as acceptable or effective for a given problem. Likewise, if they do not understand how to implement the program as intended by the program developers, even the most effective of programs will fail to have an impact on students.

There is some evidence that acceptability influences teachers' implementation of programs. Tanol (2010) applied the concepts of treatment acceptability to the prediction and explanation of variations in treatment fidelity in a year-long classroom intervention. Tanol reported moderate relationships ($r = .28$ to $.53$) between teacher's ratings of acceptability and their fidelity of implementation across time points within the intervention. Similarly, Allinder and Oats (1997) reported that teachers who showed higher acceptability ratings for a curriculum-based measure program used significantly more assessment probes and set higher goals than those with low acceptability ratings (with effect sizes greater than 1 SD). In one of the few studies to measure both acceptability and understanding, Kurita and Zarbatany (1991) measured acceptability, time efficiency, and familiarity (which could be considered a self-reported measure of understanding) with six classroom strategies for raising student motivation and reported strong correlations between these measures ($r = .41$ to $.89$).

Timing of measurement of acceptability has been found to be important as well. Peterson and McConnell (1996) reported weak correlations with overall acceptability and treatment integrity, but measured acceptability pre-treatment, which Tanol (2010) hypothesized may have reduced the correlations. Similarly, Gresham (2009) speculated that acceptability measured after a treatment has begun will be more accurate because of the greater experience teachers have with the intervention. Peterson and McConnell (1996) also noted that acceptability of the treatments was quite high, perhaps indicating restriction of range that might reduce correlations with other variables (indeed, the only significant positive correlations observed with treatment integrity were the items that had the lowest average rating for acceptability).

2. Measures of components of treatment fidelity

The choice of measures of treatment fidelity and its predictors are critical. In particular, self-report measures of treatment fidelity and related factors like acceptability are problematic because of the inherent demand characteristics (acquiescence and/or social desirability effects) such that teachers will be inclined to overestimate their use and quality of implementation and say positive things about the program (Lee, Penfield, & Maerten-Rivera, 2009; Mullens et al., 1999; Wickstrom, Jones, LaFleur, & Witt, 1998). All measures have limitations, of course. Even using observations would overestimate structural integrity measures (e.g., frequency of use) because teachers may increase their use of program components while being observed (at least in the

Download English Version:

<https://daneshyari.com/en/article/372620>

Download Persian Version:

<https://daneshyari.com/article/372620>

[Daneshyari.com](https://daneshyari.com)