Contemporary Educational Psychology 40 (2015) 112–120

Contents lists available at ScienceDirect

Contemporary Educational Psychology

journal homepage: www.elsevier.com/locate/cedpsych

Serious challenges require serious scholarship: Integrating implementation science into the scholarly discourse

Jeffrey Alan Greene *

University of North Carolina at Chapel Hill, United States

ARTICLE INFO

Article history:
Available online 22 October 2014

Keywords:
Implementation fidelity
Translational research
Educational psychology
Mixed methods
Causality

ABSTRACT

The articles in this special issue illustrate the challenges of implementing interventions in school contexts, as well as the lessons that can be learned from such work. Being responsive to the challenges and affordances of educational contexts requires studying not just the treatment, but also what happens before, during, and after implementation. Scholarship on implementation science and curricular design can be productively integrated into intervention development and research, as the authors in this special issue have shown. Gathering data regarding the entirety of the intervention implementation can result in powerful lessons for the field, but researchers, and journal editors, must apply the same standards of rigor for reporting implementation fidelity as they do for reporting psychometrics or statistical analyses. As this special issue shows, when the field of educational psychology takes seriously the scholarship of implementation, the result is positive implications for theory, practice, and the preparation of future scholars.

1. Introduction

Quite simply, over the last decade, doing research in schools has gotten both harder and more important, as educators struggle to do their work with fewer resources, and researchers likewise wrestle with how to intervene in ethical, benevolent, and congruent ways within that context (Dinella, 2009; O'Connell & Gray, 2011). Therefore, the focus of this special issue is timely, in that educational psychology as a field is moving toward more explicit recognition, investigation, and reconciliation of the "dynamic complexities of spending extended research time in classrooms" (Murphy & Cromley, 2013, p. 107). The challenge of translating research to practice extends well beyond the very real difficulties of finding schools to work with, or developing materials for teachers, and into substantive theoretical issues including how to take what has worked in the past and situate it in new contexts. As more and more educational psychologists explore Pasteur's Quadrant (i.e., applied, innovative research; Stokes, 1997), these dynamic complexities, and the "tricks of the trade" about how to understand and manage them, which used to be discussed in the hallways at conferences, must instead be brought fully into the light of the scholarly discourse. It will take the field's collective knowledge and skills to identify the toolbox of techniques needed when researchers seriously wrestle with "situating relevant interventions within existing curricula, dealing with varying student abilities, school cultures, classroom enclaves, pedagogical nuances, and a general malaise toward research" (Murphy & Cromley, 2013, p. 107). The growing literature on fidelity of implementation in education (e.g., O'Donnell, 2008), implementation science in mental health and educational policy research (Fixsen, Blase, Duda, Naoom, & Van Dyke, 2010), and design-based research in the learning sciences (Brown, 1992; Kelly, Lesh, & Baek, 2008), can inform how educational psychologists translate research to practice, and then to contextualization and dissemination.

Aside from the ethical and scholarly obligations that compel us to take seriously these issues, the education practitioner communities are justifiably growing less and less likely to partner with researchers who fail to understand how schools work, and how to navigate the many challenges of translating theory into practice (Dinella, 2009). Indeed, the Education Sciences Reform Act of 2002 created an entire organization, the Institute of Education Sciences (IES), dedicated not only to funding, synthesizing, and disseminating studies of educational practice and policy, but also training future researchers to do that work. One of the main foci of this organization has been careful attention to the issues of practical implementation and scale-up of evidence-based educational interventions. Educational psychologists, and the educators and students with whom we work, benefit from rigorous research into

* Address: 113 Peabody Hall, CB#3500, School of Education, Learning Sciences and Psychological Studies Program, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3500, United States. Fax: +1 919 843 2614.
E-mail address: jagreene@email.unc.edu

http://dx.doi.org/10.1016/j.cedpsych.2014.10.007
0361-476X/

not only whether treatment groups differ from control, but why, and how these interventions can be contextualized and disseminated across varying contexts and people. This special issue highlights the great challenges, and important lessons, that can be unveiled when researchers open up the "black box" of the treatment variable, and rigorously study what actually happens before, during, and after implementation.

1.1. Terminology

When researchers and educators take seriously the need to study the entire process of translating theory to practice and context, a common scholarly language develops by necessity. Studies that examine an intervention's *efficacy*, where an intervention's effects are tested in settings where researchers control much of the implementation, must be followed by studies of the intervention's *efficiency*, or the effects of the intervention when implemented by communities of practice, rather than researchers (Warren, Domitrovich, & Greenberg, 2009). In both efficacy and efficiency trials, researchers must pay close attention to the *fidelity of implementation*, also called *implementation* or *intervention fidelity*, which is the degree to which the intervention was implemented as designed (Nelson, Cordray, Hulleman, & Sommer, 2012; O'Donnell, 2008). This fidelity should be measured in multiple ways. One metric is *dosage*, which is how often, or for how long, participants (e.g., students) receive the intervention. For example, despite training and support, the implementers of an intervention may not enact it as often as the designers intended (e.g., Star et al., 2015). The effect of the treatment, then, may vary depending upon the "dosage" of the intervention each implementer administered to participants. In this special issue Rubie-Davies et al. (2015) argued that teachers with whom they worked enacted the intervention more frequently in mathematics lessons than literacy, perhaps explaining the lack of a treatment effect in the latter. Without careful measurement of dosage, the authors would be left to only speculate on possible explanations for this differential effect.

Fidelity of implementation can also be measured in terms of the *quality of the delivery* of the intervention, meaning the degree to which implementers (e.g., researchers, teachers) follow the procedures outlined for the intervention. Festas et al. (2015) found that their treatment teachers implemented 78% of the intervention activities, on average. This lack of complete implementation leads to an important question: which components were not covered, and did that affect the quality, not just the intensity, of the treatment? Cervetti, Kulikowich, and Bravo (2015) found significant variation in the degree to which teachers utilized the intervention strategies, and this variation seemed related to student outcomes.

Finally, the *quality of the implementer training* can also be investigated, including whether there was sufficient initial and on-going support throughout the intervention to ensure that trainees have acquired the necessary knowledge, skills, or attitudes to implement the treatment as designed (e.g., Festas et al., 2015; Harris, Graham, & Adkins, 2015; McMaster et al., 2015). These three aspects of implementation fidelity should be measured using multiple indicators including self-report surveys, interviews, and observations, among other methods (Nelson et al., 2012). Careful measurement and analysis of fidelity of implementation data are essential to determining whether a lack of statistically and/or practically significant findings is due to *theory failure*, i.e., flaws in the ideas that informed the design of the intervention, or *implementation failure* (e.g., Star et al., 2015).

Efficiency trials, with careful examination of implementation fidelity, are particularly necessary before initiating wide-scale *dissemination* of the intervention to practitioners beyond those who have direct contact with the researchers. Practitioners often have to modify the intervention to fit local contextual issues, but this *implementation drift* is not necessarily a bad thing. Indeed, often these adaptations are necessary just to ensure that the local community accepts the intervention. Such adaptations require, however, that researchers identify, and then communicate, the *active ingredients* of the intervention, or those parts that seem essential to deriving effects across contexts, and that therefore should not be varied. Everything else in an intervention can and often will be adapted by practitioners, by necessity. Studies of implementation fidelity, within efficacy and efficiency trials, can help researchers identify these active ingredients (Warren et al., 2009).

1.2. Overview of this response article

Overall, school-based research requires that researchers be responsive, adaptive, and creative when confronted with the dynamic complexities of schools, as well as systematic in documenting and studying these efforts (Dinella, 2009; Nelson et al., 2012). The results of these investigations can inform how theory is refined as well as how it is translated to practice and context. Such results can also be used to improve the methods for preparing implementers to enact the intervention. Multiple investigations of the same treatment can provide information regarding the different ways future researchers can construct their work to have the greatest likelihood of being successful. Publishing these findings in the scholarly literature is the first step toward more fully integrating the lessons, like those learned in this special issue, into the knowledge base that future researchers learn in our preparation programs.

2. Review of the special issue

2.1. Measuring and communicating implementation fidelity

For education researchers who aspire to translate their work to multiple contexts, a broad-bandwidth approach to data collection is required from the very outset of design conceptualization. Researchers must explore questions beyond whether participants in conditions differed on a single outcome measure. True understanding of an intervention, including its active ingredients and its likelihood of scalability, requires investigating the entirety of the participants' (i.e., students, but also teachers, administrators, and other school partners) experience (Nelson et al., 2012). The canonical "treatment variable" really only captures the researchers' intent-to-treat effect, via random assignment. What actually happens during the intervention, including who was actually treated, what the treatment was, and how it was experienced, is obscured by the "black box" of the treatment variable. Indeed, it can be easy to overlook that, in the two studies of the SRSD in this issue (Festas et al., 2015; Harris et al., 2015), the treatment variable was actually problem-based professional development (PBPD). The SRSD aspects of the study were endogenous, meaning that the researchers did not directly manipulate them; rather they were indirectly manipulated via the PBPD treatment. Given this, it was not surprising that SRSD was implemented with varying degrees of fidelity by the implementers (i.e., the teachers in this study). By carefully assessing implementation fidelity (O'Donnell, 2008), the authors of these studies could capture both the treatment effect (i.e., PBPD exposure) as well as how teachers then implemented SRSD with their students. Both "variables" are needed to understand any student effects.

Therefore, participants' experiences during the intervention are critical data that must be captured, analyzed, and integrated into inferences about the work. McMaster et al. (2015) utilized a wide