Abstract Title Page
Not included in page count.

Title: Applying the Five-Step Model of Fidelity Assessment to a Randomized Experiment of a High School STEM Intervention

Authors and Affiliations: Jason P. Kopp and Dr. Chris S. Hulleman, James Madison University
Dr. Judith M. Harackiewicz and Chris Rozek, University of Wisconsin-Madison

Abstract Body

Limit 4 pages single-spaced.

Background / Context:

Description of prior research and its intellectual context.

Assessing fidelity of implementation is becoming increasingly important in education research, in particular as a tool for understanding variations in treatment effectiveness. Fidelity of implementation is defined as "the determination of how well an intervention is implemented in comparison with the original program design during an efficacy and/or effectiveness study" (O'Donnell, 2008). Failing to achieve a high degree of implementation fidelity can lead to wide discrepancies between intended and achieved treatment strength (Hulleman & Cordray, 2009). Accordingly, it has been shown that fidelity can have significant effects on outcomes (e.g., McIntyre, Gresham, DiGennaro, & Reed, 2007). Despite the importance of fidelity, there has been a lack of a unifying framework for fidelity assessment across various disciplines (Durlak & DuPre, 2008). To this end, Nelson, Cordray, Hulleman, Sommer, & Darrow (under review) created a simple five-step procedure for assessing intervention fidelity. Their procedure involves 1) specifying the intervention model, 2) identifying appropriate fidelity measures, 3) determining the reliability and validity of those measures, 4) combining fidelity indices, if appropriate, and 5) linking fidelity and outcome measures. Despite the theoretical promise of this model, it has yet to be applied to actual interventions and outcomes.

To address this gap in the literature, we applied the above five-step procedure to data from a randomized field intervention designed to increase enrollment and motivation in STEM courses (science, technology, engineering, and mathematics) for high school teens (Harackiewicz, Rozek, Hulleman, & Hyde, under review). The intervention was based on expectancy-value theory (Eccles & Wigfield, 2002), and aimed to increase the perceived utility value (i.e., usefulness and relevance) of STEM courses for both parents and teens.

Purpose / Objective / Research Question / Focus of Study:

Description of the focus of the research.

To apply the Nelson et al. (under review) five-step model to the randomized field experiment conducted by Harackiewicz et al. (under review).

Setting:

Description of the research location.

The intervention took place with teens from 108 different high schools in the Midwest, over the course of a 15-month period during 10^{th} and 11^{th} grade.

Population / Participants / Subjects:

Description of the participants in the study: who, how many, key features, or characteristics.

Data were collected from 188 teens (88 girls, 100 boys) and parents from the longitudinal Wisconsin Study of Families and Work (Hyde, Klein, Essex, & Clark, 1995). The sample was 90% Caucasian, 2% African American, 1% Native American, and 7% biracial or multiracial. The sample of families had been previously recruited for a study of maternity leave and health (before the birth of the teens), and had been followed longitudinally through the teens' final year of high school. Families were randomly assigned to either the treatment or control condition.

Intervention / Program / Practice:

Description of the intervention, program, or practice, including details of administration and duration.

The goal of this intervention was to encourage parents to talk to their teens about the utility value of STEM courses for their futures, which would then lead to increased STEM course-taking in high school. To this end, parents of teens in the treatment condition received a multi-faceted intervention over the course of a 15-month period during 10th and 11th grade. The intervention consisted of three primary components:

1) In October of 10th grade, researchers mailed households a brochure titled "Making Connections: Helping Your Teen Find Value in School," addressed to parents. This brochure contained information about the value and importance of STEM courses, as well as guidance on having conversations with teens regarding this topic.
2) In January of 11th grade, researchers mailed households a brochure titled "Making Connections: Helping Your Teen with the Choices Ahead" separately to each parent. This brochure also emphasized the importance of science and mathematics and provided guidance regarding having conversations with teens, but utilized different and more specific examples. This brochure included a link to a password-protected website, "Choices Ahead." The website included a wealth of resources regarding STEM fields and the option to send a link to the website to their teen.
3) In the spring of 11th grade, parents were asked to complete an online questionnaire evaluating the "Choices Ahead" website. This was intended to increase exposure to, and interaction with, the website.

Significance / Novelty of study:

Description of what is missing in previous work and the contribution the study makes.

Although fidelity assessment has become an increasingly important component of the evaluation of educational interventions, there has been little systematic analysis of the extent of fidelity assessment and its role in understanding variations in treatment effectiveness. This study applies a five-step model of fidelity assessment to data collected as part of a randomized field experiment of an intervention intended to increase STEM course-taking in high school.

Statistical, Measurement, or Econometric Model:

Description of the proposed new methods or novel applications of existing methods.

In this proposal, we apply the five-step model of fidelity measurement to data collected as part of the Harackiewicz et al. (under review) field experiment. Here, we outline how the model applies to the data. At the conference, we will present analyses incorporating the fidelity indices to understand variations in treatment effectiveness.

Step 1: Specifying the intervention model. Nelson et al. (under review) suggest specifying two logic models: the change model and the operational model (Knowlton & Phillips, 2009). The change model includes a network of causal relations between constructs included in the intervention. This allows the researchers to delineate exactly how and why the intervention should work (see Figure 1). All of the intervention materials (two brochures and one website) contained two distinct intervention pieces that were hypothesized to create changes in the parents. First, all materials included information about the importance of STEM. Second, the materials provided instruction on having a conversation with the teen regarding the utility of STEM. These two pieces should combine to increase conversations between the teen and parents on STEM utility. We hypothesized that this would then increase teen STEM course-taking in the last two years of high school.

After specifying the change model, Nelson et al. (under review) recommend elaborating on the change logic model by specifying an operational logic model. The operational model builds upon the change model, and serves as the basis for fidelity assessment by specifying the materials, resources, and personnel required to implement the intervention. As presented in Figure 2, the two components of the intervention (information on the utility of STEM and guidance on conversations with teen) are operationalized as information contained in the two brochures and the website (see Figure 2). Parents must read and process the information in order to have increased utility value for STEM and conversations with their teen. The conversations themselves must be of sufficient frequency and quality to facilitate behavior change in the teen. Parents may also share materials with their teens. Finally, increased STEM course-taking will be evidenced on teens' high school transcripts.

Step 2: Identifying appropriate fidelity measures. Once the operational model is specified, appropriate fidelity indices can be selected that directly map to components of the operational model. The first two components of the operational model (the sending and receipt of the materials, and the extent to which parents read and processed the information) were assessed via interviews with the parents and number of logins to the website. The extent to which conversations took place between parent and teen, the quality of those conversations, and the amount of materials shared with the teen were all assessed via interviews and self-report surveys with both parents and teens. Collectively, this information enables us to determine the extent to which the individual components of the operational model were implemented with fidelity. The interview data is the richest and most abundant fidelity data, and is currently being coded into quantitative variables. At the conference, we will present analyses of these variables.

Step 3: Determine index reliability and validity. Before incorporating fidelity information within the analysis of treatment variation, reliability and validity evidence must be gathered for the different fidelity measures. Reliability involves the extent to which the fidelity measures inter-correlate. For instance, we would expect that parent and teen reports of conversations would be consistent within household. Validity involves the extent to which the fidelity instruments measure what is intended (i.e., construct validity). For instance, assessing the frequency of conversations with teens, but not the quality, may fail to cover the breadth of the fidelity construct (Shadish, Cook, & Campbell, 2002).

Step 4: Combining indices. If several fidelity indices are highly interrelated, they can be combined into composite fidelity indices. Composite fidelity indices are useful, as they can be more easily utilized to determine achieved relative strength of the intervention (Hulleman & Cordray, 2009). These indices will be presented at the conference.

Step 5: Linking fidelity measures to outcome measures. Once the appropriate fidelity measures have been combined, fidelity can be linked to outcome measures. It is possible that fidelity related to some components is more strongly linked to outcomes than others. This can help identify core vs. ancillary intervention components. For instance, we may find that the frequency of conversations with teens is strongly related to enrollment in STEM courses, but measures of conversational quality are not related to enrollment. This can shape future intervention applications and research. At the conference we will present analyses which calculate indices of achieved relative strength (Hulleman & Cordray, 2009) and incorporate the fidelity indices into analyses of treatment effects using techniques outlined by Bloom (2005), Peck (2003), and Schochet and Burghard (2007).

Usefulness / Applicability of Method:

Demonstration of the usefulness of the proposed methods using hypothetical or real data.

The complete five-step method is demonstrated using existing data from a randomized field experiment of an intervention designed to increase STEM course-taking in high school. Prior demonstrations of the five-step model have used examples that highlight only one or two aspects of the model (e.g., Cordray & Hulleman, 2009; Hulleman, 2011).

Research Design:

Description of the research design (e.g., qualitative case study, quasi-experimental design, secondary analysis, analytic essay, randomized field trial).

The research presented in this paper is the analysis of intervention fidelity data from a randomized experiment of 188 teens (100 in the experimental group, 88 in control). The primary dependent variable was enrollment in STEM courses in 11th and 12th grade. As previously mentioned, the design of the current study is to assess the fidelity of the above study utilizing the five-step procedure outlined by Nelson and colleagues (under review).

Data Collection and Analysis:

Description of the methods for collecting and analyzing data.

Transcripts were collected for the 181 teens (we were unable to access transcripts from seven teens) participating in the study to determine STEM course enrollment. Teens and parents were asked to complete a survey in the 12th grade. Surveys were obtained from 171 teens, 169 mothers, and 126 fathers. Teens and parents were also interviewed regarding their interaction with the treatment materials in both 10th and 11th grade, as well the frequency and quality of conversations about STEM utility. As part of Steps 3, 4, and 5 of fidelity assessment, these fidelity measures will be used to create indices of achieved relative strength and to predict outcomes. These results will be presented at the conference.

Findings / Results:

Description of the main findings with specific details.

With regards to the primary outcome variable, students in the treatment group took significantly more STEM courses in their last two years of high school (M = 8.31 semesters) than students in the control group (M = 7.50) (Harackiewicz et al., under review). The data on the fidelity measures is currently being coded and analyzed. As part of Step 5 of the fidelity assessment process, indices of achieved relative strength will be created, and the extent to which fidelity helps explains variations in the treatment effect will be calculated.

Conclusions:

Description of conclusions, recommendations, and limitations based on findings.

Fidelity assessment holds great promise for the field of education research. We hope that this study will show the utility of the 5-step model proposed by Nelson et al. for providing an effective framework for fidelity assessment.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

Bloom, H. (2005). *Learning more from social experiments: Evolving analytic approaches*. Russell Sage Foundation.

Cordray, D. S., & Hulleman, C. S. (2009, June). *Assessing intervention fidelity in RCTs: Models, methods and modes of analysis*. Invited panel session at the 2009 Institute of Education Sciences Research Conference, Washington, D.C.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41,* 327-350.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53,* 109-132.

Harackiewicz, J. M., Rozek, C. S., Hulleman, C. S., & Hyde, J. S. (under review). Helping parents motivate their teens in mathematics and science: An experimental test.

Hulleman, C. S. (2011, September). *Conceptualizing intervention fidelity: Implications for measurement, design, and analysis*. Paper presentation at the Fall Conference of the Society for Research on Educational Effectiveness, Washington, DC.

Hulleman, C. S., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness, 2,* 88-110.

Hyde, J. S., Klein, M. H., Essex, M. J., & Clark, R. (1995). Maternity leave and women's mental health. *Psychology of Women Quarterly, 19,* 257-285.

Knowlton, L. W., & Phillips, C. C. (2009). *The logic model guidebook: Better strategies for great results.* Washington, D. C.: Sage.

McIntyre, L. L., Gresham, F. M., DiGennaro, F. D., & Reed, D. D. (2007). Treatment integrity of school-based interventions with children in the Journal of Applied Behavior Analysis, 1991-2005. *Journal of Applied Behavior Analysis, 40,* 659-672.

Nelson, M. C., Cordray, D. S., Hulleman, C. S., Sommer, E. C., & Darrow, C. L. (under review). A procedure for assessing intervention fidelity in experiments testing educational and behavioral interventions.

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research, 78,* 33-84.

Peck, L. R. (2003). Subgroup analysis in social experiments: Measuring program impacts based on post-treatment choice. *American Journal of Evaluation, 2(24)*, 157-187.

Schochet, P. Z., & Burghard, J. (2007). Using propensity scoring to estimate program-related subgroup impacts in experimental program evaluation. *Evaluation Review, 31*, 95-120.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin Company.

Appendix B. Tables and Figures

Not included in page count.

Figure 1: Harackiewicz et al. (under review) change logic model

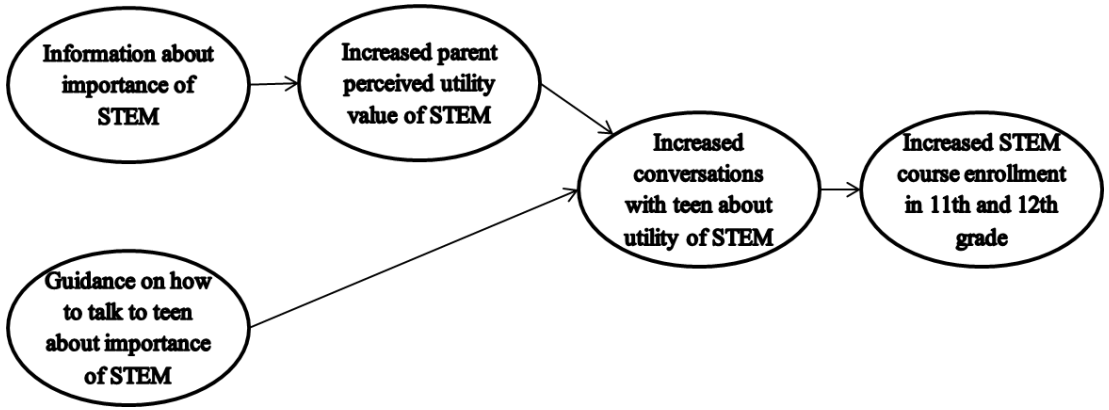

Figure 2: Harackiewicz et al. (under review) operational logic model