Reviewer's report

Title: Blueprints for Violence Prevention Process Evaluation of the Life Skills Training Program: Factors Related to Implementation Fidelity

Version: Date: 18 May 2006

Reviewer: Dean Fixsen

Reviewer's report:

General
Review of Mahalic, Fagan, Argamaso MS

Reviewer: Dean Fixsen and Frances Wallace

The MS "Blueprints for violence prevention process evaluation of the Life Skills Training program: factors related to implementation fidelity" by Mihalic, Fagan, & Argamaso was a pleasure to read. Mihalic and colleagues are well known researchers working out of the famous Center for the Study and Prevention of Violence (CSPV). The study described in the MS is focused on the attempted implementation of the Life Skills Training (LST) Program in 404 schools located at 105 sites that touched the lives of 172,285 students. The scope alone makes this a remarkable implementation research project and provides ample opportunities to realize the promise of the title: "factors related to implementation fidelity."

It is not news to contributors to and readers of the Journal of Implementation Science that implementation efforts are complex and fraught with many issues at every turn. No one study nor even a research career of studies will tease out all we need to know about "what works" best for implementation. We are left with the time worn device of trying to milk all we can from each study so we can generate better hypotheses and measures as we go along the path toward evidence-based implementation methods. This is a good thing and a device well-used by these authors as they describe the results of their quantitative and qualitative assessments of the study they carried out over many years.

In studies of implementation, the independent variables are the specific methods used to attempt to help others make use of a well-described and well-evaluated program. While reading the MS, we tried to outline the methods employed by CSPV and its partner organization, National Health Promotion Associates (NHPA), described as "providers of the LST program." On the LST website, we are informed that, "Training services are available to you through National Health Promotion Associates, Inc. (NHPA), a health consulting, research, and development firm developed by Dr. Botvin, the developer of the Life Skills Training program." Thus, it seems that NHPA is the approved purveyor for the LST program, a good thing when attempting to implement an evidence-based program.

The independent variables (implementation methods)

Site selection: An OJJDP RFP with CSPV as the agency to decide who was funded; not many schools (or unidentified others) applied; written commitment from school officials; feasibility visits conducted by CSPV and NHPA personnel to discuss the LST program, explain requirements, assess commitment, and discuss issues. Given the paucity of applicants, only a few applicants were excluded based on lack of support from administrators/teachers or unable to allow required observations in the classroom (the two reasons cited in the text).

More information is needed re: the RFP process, amounts of the awards, what the awards paid for, etc., all of which may influence the interpretation of the information from this study. Also, given the importance of the exploration stage of implementation (re: giving and getting accurate information, establishing the need for a program, identifying and supporting champions, identifying and supporting leaders, securing buy in, etc), more information is needed. What was done to "present the core elements," "explain requirements," "address concerns," and "address concerns"? Who was involved? What was the level of knowledge and expertise of the presenters? Who were the audiences for these messages? What protocols were followed? What variations were needed across the 404 schools? What is a site? How are sites and schools connected? Were sites and schools approached independently or differently?

Given the breadth of activity, the authors could make a substantial contribution to the literature on the initial
stages of implementation. It is curious that no demographic information is presented about the teachers, schools, and sites who are the "subjects" of this study of implementation. In addition, no data were collected about the behavior of the purveyors (CSPV, NHPA, and company) or the impacts on buy in and support from key stakeholders (see Panzano, et al, in press for a good example of data collection protocols for the exploration stage). Given that schools seemed to have issues with buy-in and maintaining the program, information on the exploration process would provide more insight into the implementation of the program.

Initial implementation: "Each site received a two-day training workshop;" LST instructors and evaluators hired to do classroom observations had to attend but training was optional for local coordinators and school administrators; compliance with the LST curriculum and instructional techniques was assessed in about 1/3 of the classes by the evaluators hire for this project; technical assistance was provided by CSPV and NHPA via annual site visits (interviews and discussions with local coordinators, principals, evaluators, and "some teachers"); schools could request more technical assistance.

Much more information is needed re: what the purveyors did during the initial implementation stage and with whom they did it. Fundamental information is missing. What was the content of training (an outline of topics) and how was the information presented (didactic, behavior rehearsal). How many Level 1, 2, and 3 workshops actually were presented in the three-year period? Who exactly was trained at each site for each set of workshops? We are never told the number of teachers (or other instructors) or class sizes -- was it just one instructor for all the students at each site? Apparently, some instructors were not teachers and some sites were not schools -- what were they and how many and why? How many of each kind were scheduled for training, how many showed up, how many finished training (were there criteria for mastery during training), and how many actually began using the LST curriculum (separate data by year for Level 1 training, Level 2 training, and Level 3 training)? "Staff turnover after training was common" -- how many teachers at how many sites each year? How many students did not receive LST as a result of turnover, missed training, refusals, etc, etc? The authors only say that 100% of exposure to the program was not obtained (which was one of the identified research questions). Given the potentially critical on-site coaching role of the local coordinators, why was their attendance at training optional and not required? How many local coordinators attended training? Was all of the preparation work done for all 404 schools at once? How was the purveyor activities scheduled to get all this done? In the conclusion section the authors indicate that some untrained teachers delivered the program. How many untrained teachers delivered the curriculum? Were these teachers included in the final outcome results (the 70% completing surveys)?

The authors mention that additional violence prevention lessons were added to the LST curriculum and that these lessons were an option. However, the authors do not tell us the nature of these additional lessons or how many schools chose to add the lessons or if the addition of these lessons had any impact on the implementation of the intervention (was there less support for the programs that chose to use the additional lessons? More student participation? Is this a confounding or mediating factor?)

Successful implementation seems to depend on the comprehensiveness and depth of training, coaching, and support provided to practitioners (e.g. LST instructors) by program purveyors. For an implementation project of this scale, the reader will be very interested in how the purveyors went about the business of helping lots of instructors in over 400 schools learn the LST program (in three stages) and support their use of the program with students. More detailed descriptions of purveyor activities can be tied to descriptive data as suggested above to give the reader a better picture of how this process unfolded on this scale. Again, no data were collected about the behavior of the purveyors or the impacts on teachers, school (and other) administrators, and so on during the initial implementation stage.

Measures: The data sources are self reports of one kind or another and the questions asked are sparse and generally non-specific. For example, "LST training quality" consists of a rating by the local coordinators at the end of the three-year grant (1 = a barrier to implementation, 5 = an asset to implementation) coupled with a rating by the teachers and a rating by the trainers themselves (from poor to excellent). These questions are too general to be very helpful to anyone seeking to improve the quality of training. In addition, teachers were rating the training after the workshop without having any experience in implementing the program. Thus, their ratings as to whether the training was a barrier or a facilitator would be a prediction on the part of the teachers and not a reflection of their experience with the curriculum. Essentially the ratings of the training immediately after the workshop would be rating the experience of the training session and not the quality and utility of the training in regard to implementation. Others have asked questions about specific aspects of training content (e.g. clarity of the information), training methods (e.g. usefulness of behavior rehearsals), and training outcomes (e.g. self efficacy). High or low scores then help to inform trainers and others about changes that are needed to improve training benefits. "Administrative support" sounds like an important dimension but consists of one question asked of teachers (not at all supportive to very supportive). In addition, there is very little information about the protocols that were followed to collect these
data -- when, where, with whom, by whom, etc. We suspect the general lack of effects found in this study might be due to the inadequate measures of the variables under study (not much room for variability in a single question scored on 3-point scale).

In addition, very little information is provided re: the major outcome variables (implementation score = evaluator observation of teachers' compliance with the LST curriculum), dosage (teacher report at the end of each year of his/her use of all lessons), quality of delivery (evaluator observation of teachers' "frequency of use" of the LST teaching techniques -- modeling, demonstration, and behavior rehearsal), and participant responsiveness (teacher report annually of the percent of students who participated in LST activities the teacher taught). Again, very sparse and general information.

The teacher adherence results indicate the range for individual teachers was from 0-100%. However, the measure of fidelity used in the paper employs the very unusual method of summarizing this information by site (not by teacher) across all three years (not by class) to arrive at a score for each site to enter into the analysis. Intervention fidelity typically is measured at the practitioner level (e.g. an MST Therapist) with regard to a specific intervention (teacher behavior in a series of classes; an MST series of family visits) to assess the proximate relationship between practitioner behavior (more or less in compliance with the intervention protocols) and consumer outcomes (more or less benefit to students). Fidelity is seen as a proxy for outcomes (high fidelity = better outcomes, low fidelity = poor outcomes). To be useful, the relationship between the measure of fidelity and the measures of outcomes needs to be demonstrated through research. The authors mention the use of the fidelity checklist used for the LST research trials, however, no information is provided on the checklist (whatâ€™s included, reliability, validity, etc.). We looked at the LST website but could find no reference or data on fidelity to help us understand this better. In any case, the outcome measures and the methods used to collect and average them require considerably more explanation to help the reader understand what was done and why is was done that way.

Throughout the manuscript the authors use intervention fidelity and implementation fidelity interchangeably. The LST fidelity checklist is a measure of intervention fidelity and, ideally, reflects the core components of the intervention as used by teachers. On the other hand, implementation fidelity would provide a measure of the core components of implementation activities used by purveyors. Thus, "implementation fidelity," the subtitle for this paper, implies a whole different set of measures than those found in this study. Intervention fidelity (was LST taught in compliance with the LST curriculum, were the required teaching methods used, did the teachers display competence in how they taught the materials) concerns the practitioners' behavior. Implementation fidelity is a different thing (was the site selection protocol used as prescribed, did training content conform to the training protocol, did pre-post tests show that the trainees actually learned the required LST content, did the local coordinators provide support and assistance to the teachers as required by the protocol, etc) and concerns the behavior of the purveyors.

In summary, this is a fascinating attempt to implement an evidence-based program on a large scale. The data sources and questions are too general and the questions to sample the important dimensions are too few to allow much in the way of meaningful data analysis. However, this is an interesting case study of attempting to implement the LST program in over 400 schools. More information on the processes used coupled with descriptive data associated with those processes would provide very useful information to the reader.

Reference
Panzano, P. C., Seffrin, B., Chaney-Jones, S., Roth, D., Crane-Ross, D., Massatti, R., et al. (in press). The innovation diffusion and adoption research project (IDARP): Moving from the diffusion of research results to promoting the adoption of evidence-based innovations in the Ohio mental health system. New Research in Mental Health, 16.
Discretionary Revisions (which the author can choose to ignore)

What next?: Reject as not sufficiently sound

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.