Reviewer's report

Title: Factors influencing sustained implementation and fidelity thereof and related student outcomes of an evidence-based national HIV prevention program

Version: 2  Date: 6 December 2014

Reviewer: Louise Ann Rohrbach

Reviewer's report:

Overall, this paper well written and should be of interest to implementation scientists. My suggestions regarding revisions are below.

Major Compulsory Revisions

1. The fifth paragraph in the Introduction ("Given the complex...") includes a description of studies that are not relevant to the present study. I would suggest revising or this paragraph to focus on findings from previous school-based studies. Also, I would suggest that the authors look at a recent special issue of the Journal of Adolescent Health (V. 54, S1-S2, 2014) that included several studies of implementation of pregnancy prevention programs that may be relevant to the present paper.

2. The Introduction section should establish why it is important to study implementation. Specifically, what have studies shown regarding the association between implementation and outcomes (particularly school-based prevention studies)? For example, see the Durlak and DuPre (2008) review article that is cited in the Discussion section.

3. The authors use of the term "sustained implementation" is not consistent with the literature. Usually, sustained implementation refers to implementation of a program over time, after initial implementation (i.e., maintenance or institutionalization of it in an organization). I would suggest that the amount of the intervention that teachers delivered in this study is not a measure of "sustained implementation," but rather a measure of implementation dosage. I would suggest that the authors refer to this measure in a way that is consistent with the implementation literature.

4. In the Discussion section, under the header "Teacher's implementation group membership and student outcomes," the authors state that "the quality of implementation was causally related to student outcomes." This statement is an over-interpretation of the findings. As the title of Table 4 states, the analysis examines an association between implementation and outcomes, not a causal relationship. In order to establish causality, the authors would need to either conduct a randomized design, assigning teachers to implementation groups, or conduct a mediation analysis to determine if varying levels of implementation mediated changes in student outcomes.
5. In some cases in the Results section, the results are worded as if there was a comparison between two implementation groups (e.g., High/Moderate vs. Low) rather than a comparison among three groups. For example, in the first paragraph under the subheading "Association of teachers' implementation group membership with teachers' characteristics and pre- and post-implementation perceptions" (findings regarding length of time as a teacher and island of teacher). Overall, the authors should be consistent in inclusion of the F statistic in parentheses after their statements regarding group differences, to remind readers that the statistic represents a comparison among three groups.

6. In the last paragraph of the Analysis section, the authors should clarify that in the multi-level analysis, they conducted two comparisons: high vs. low implementers and moderate vs. low implementers. As written, the difference in what is compared in the multi-level analysis relative what is compared in all other analyses in the paper (ANOVA for 3 groups) is not clear.

7. In the Discussion section, the authors should discuss the differences in findings from the multi-level analysis and the ANOVA analysis reported in Table 3, and what they mean.

Minor Essential Revisions

1. In the Introduction, fourth paragraph, I would not describe the literature that suggests adaptations are inevitable as "robust." Many authors have suggested or offered an opinion that this is true, in at least some cases based on anecdotal evidence; there is not a strong body of research on the extent and nature of adaptation of evidence-based programs when they are implemented in the "real world."

2. In the Introduction, fourth paragraph, the sentence "A particularly comprehensive examination of fidelity..." needs a citation.

3. Introduction, paragraph 6, "as it was shown to be effective" is an awkward wording. I would suggest using "as it was designed" or "as it was implemented in efficacy/effectiveness trials" or something similar.

4. Methods section, first paragraph under "Intervention" has a typo in the sentence that begins, "The intervention,..." 

5. Analysis section, second paragraph, the sentence that begins "The differences in..." I would revise this to make clear which differences are being examined.

6. Results section, first paragraph, there is a typo: "....twenty-eight (13.5%) taught more than 40 all activities...."

Discretionary Revisions

1. I would consider including a figure or schematic that shows the conceptual framework for your study. You are looking at a number of variables, and it would be helpful to show your hypotheses about their interrelationships.
2. Another limitation of the study is that it does not tell us much about differences between High and Moderate implementation groups. What are the practical implications of differences between these levels of implementation?

3. In the Discussion section, I would add that these results provide some support for the validity of self-reported measures of implementation. In the school-based literature, a few studies have called into question whether self-reported implementation measures are valuable, compared to direct observation of implementation.

**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.

**Declaration of competing interests:**

I declare that I have no competing interests.