Reviewer’s report

Title: Participatory Science and Innovation for Improved Sanitation and Hygiene: Process and Outcome Evaluation of Project SHINE, a School-based Intervention in Rural Tanzania.

Version: 1 Date: 07 Nov 2016

Reviewer: Josh Garn

Reviewer's report:

The paper is a pilot study of a school WASH intervention that took place in two secondary schools in Tanzania, with the goal to evaluate culturally and contextually relevant strategies to improve sanitation and hygiene. The study has the potential to make a meaningful contribution, although there are some methodological issues (all from the quantitative portion of the paper), which might require require further exploration and/or mention in the limitations section.

Major comments:

Sample size: In studies where interventions are implemented at the cluster-level and studies that are likely to have correlated/clustered data the number of clusters really matters and 2 schools is a really small sample size. I realize this is a pilot study, but the results should be presented in a way that it always keeps in context the sample size limitations. Please be aware to not overstate your quantitative findings. "Intervention effects" or "strong effect" is strong causal language for a study design that isn't well equipped for assessing causality (e.g., non-randomized allocation of intervention and no true control group, no control variables included in analyses). The small numbers of clusters (even though it is to be expected in a pilot study) should be mentioned as a limitation.

Accounting for correlated data: Differences in effects (e.g., between schools, sexes and age groups) were investigated using Chi-square and t-tests, but these analyses didn't account for clustering within schools (which I assume could have been done using linear regression with either a school-level random effect or alternatively by using a GEE model) or for the correlation over time (which was not possible, at least at the individual-level, as the authors already explain). The authors state that a "between subjects design was used where each measurement wave was treated independently (independent sample t-test)." However, this type of t-test is designed to compare two groups that have no relationship, and because the data are correlated, this analysis doesn't seem appropriate for these data. Because clustering was ignored, and it can have a substantial impact on the variance of the estimates, I think the following sentence is an overstatement: "The statistical power to find intervention effects will therefore be considerably smaller and the current analysis should therefore be regarded as conservative [24]." The authors...
might consider an analysis that further explores the accounting for the correlated data, or minimally list this as a limitation.

Accounting for confounding: The current quantitative analyses are unadjusted and don’t attempt to account for potential confounders. Because this study does not have a true control group, lack of control for potential confounders should minimally be mentioned as a limitation, but if adequate data are available the authors could explore controlling for potential confounders (e.g., one could use use an additional multiple linear regression model instead of paired t-tests).

Results from table 4 may be statistically significant but are they really meaningfully different? They might be, but we can’t actually tell because they are based on a continuous Likert scale, and we don’t have much of a frame of reference for what these Likert values mean in real life. For a couple of key variables, do you have the ability to additionally report how the proportion of a key sanitation/hygiene variable changed from baseline to endline? For example, do you have and can you report the proportion who reported using a latrine both pre and post so we can have a tangible frame of reference of how meaningful any differences are? If you can show in some way that the differences are meaningful, and not just statistical, then the results would be more compelling.

How many pupils answered baseline and follow-up surveys? I didn't see this information in the methods (I saw the baseline number in table 3 but couldn't find follow-up numbers, and couldn't find either in the actual text). Also, in table 3 the N is different in the table heading, than in the column heading (there are 140 more people, or 15% more, in the table heading than the column heading, yet the footnote only says that there was < 1% missing)

In the "Project outcomes" and other sections of the discussion, it seems like one alternative that isn't objectively mentioned, is the possibility that that the intervention may simply not have had the expected impact for some of these outcomes. The case for SHINE was often argued, but in some cases one could argue that any inherent biases may have made SHINE look better than it actually was. For example, with all of these self-reported outcomes, social desirability might have been even more likely to lead to overreporting at follow-up (rather than at baseline). The subjectivity of these self-reported measures should be noted as a limitation.

Minor comments:

There were a couple of minor errors. Please proof read for grammar, punctuation, and try to shorten run-on sentences. Spellings were not consistent throughout the manuscript, for example, sometimes the British spelling (behaviour) is used and sometimes the American spelling (behavior) is used.
Could you please explicitly define pastoralist in this context as early as possible in the manuscript, and also on first mention of the word pastoralist please move up the text on why it matters to be a pastoralist (i.e., given that pastoralists, who are semi-nomadic and typically reside in dry arid areas that experience drought and lack of access to other resources and services, are among the most marginalized populations in the world).

Objective 2 listed in paragraph beginning in line 74 makes it sound like the paper will be assessing health outcomes, and it doesn't assess health. Please clarify the wording.

I would prefer to see the p value recorded consistently, even if not significant. Statistical significance is really just based on an abritrary cut point, and I would want to see a value of 0.06 in the table just as much as a 0.04.

**Are the methods appropriate and well described?**
If not, please specify what is required in your comments to the authors.

No

**Does the work include the necessary controls?**
If not, please specify which controls are required in your comments to the authors.

No

**Are the conclusions drawn adequately supported by the data shown?**
If not, please explain in your comments to the authors.

Yes

**Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?**
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

**Quality of written English**
Please indicate the quality of language in the manuscript:

Acceptable

**Declaration of competing interests**
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal