Reviewer's report

Title: Beyond the Frontiers of Effectiveness: A Framework and a Measurement Instrument for Sustainability of Work Practices

Version: 1 Date: 15 December 2010

Reviewer: Greta Cummings

Reviewer's report:

Thank you for the opportunity to review and provide feedback on this manuscript. This manuscript describes the psychometric assessment and properties of a newly developed instrument to measure sustainability of work practices introduced through quality improvement initiatives in healthcare.

This is a very interesting and relevant topic for healthcare and health services research today.

Over all the manuscript is very well written and clearly provides the readers with the authors’ approach, thinking and related expectations. The work is founded on a theoretical framework that is described in sufficient detail, strengthening the work on developing the relevant measure. Reported are the results of SEM to test the measurement model. The authors indicate that they completed PCA however did not report the results.

• Major Compulsory Revisions

1. The authors need to make further distinctions between validity and reliability and be explicit about the forms of validity that they examined. For example, they conducted face validity with a team of experts and their research team but were not explicit about who these experts were. Were they from the same disciplines as those completing the surveys? Also on page 21, last paragraph, validity seems to be integrated with reliability. Also, the wording of some of the items is also very complex which may be an issue of validity related to the understanding of particular participants.

2. Sample size and characteristics – The sample size is reported, but the characteristics and demographics is not sufficiently described. This is particularly important when the sample is not homogenous and subgroups may be reporting differently based on their disciplinary backgrounds, knowledge and cultures.

3. Unit of analysis – the authors are silent on the expected unit of analysis which assumes that it is all at the individual level. Yet, the description and purposive sampling to reflect teams suggesting that the unit of analysis may be team, better be at the team level. With many of the questions using “we”, it suggests that the respondents are expected to respond related to their team not themselves. ICC calculations should be done to test for this, in addition to the theoretical examination of the level that each question is tapping in to.

3. SEM testing – The authors reported several fit indices including the
chi-square. The significance of the chi-square should also be reported and discussed.

4. While I agree with reporting the results of SEM testing, rather than PFA, I expected to see inter-item correlations, and the change in alpha for sequential deletion of items, as part of the psychometric assessment.

5. Presentation of the results and tables. The results section is structurally confusing in the explanation of the modeling and reduction in items. A separate table for the final model items would be advisable.

• Minor Essential Revisions

• Discretionary Revisions

**Level of interest:** An article whose findings are important to those with closely related research interests

**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