Reviewer's report

**Title:** Developing a Tool to Measure Health Worker Motivation in District Hospitals in Kenya

**Version:** 1  **Date:** 2 October 2008

**Reviewer:** Guy Kegels

**Reviewer's report:**

Review

Developing a Tool to Measure Health Worker Motivation in District Hospitals in Kenya.

Authors: Mbhandyo P, Blaauw D, Gilson L, English M.

This is a very carefully worded paper, submitted by highly respected authors. The reason why I submit this review a few days past the deadline – for which I apologise – is not only poor time management on my part, but also the fact that I wanted to allow myself some time to re-examine my initial perception and to check some more of the literature on the subject of motivation assessment, a reputedly complex issue. After careful consideration and repeated readings of the manuscript I think my conclusions should be formulated as follows.

The major point the authors can convincingly make is that the 23-questions Self-Administered-Questionnaire (SAQ) that they initially used, can be transformed into a more parsimonious SAQ with only 10 questions, organised around 3 latent factors.

The major problem I have with the paper is that, despite very careful wording, its main purpose seems to be that this tool can result in a single score summarising a hospital's workforce motivation. The careful wording shows in formulations like ‘the score could be an important component of motivation analysis’ (p.8); ‘hospital mean scores should be interpreted cautiously’ (p.8); ‘interpreting the summary scores as typical of Kenyan hospital settings should be approached cautiously’ (p.11); ‘we have been able to develop a potentially rapid motivation measurement tool…’ (p.11), and several other similar turns of phrase (my emphases added).

The reasons why this apparent purpose to present a composite score for motivation as a tool to measure motivation is problematic for me, are several.

First, although it can be understandable in terms of research elegance to develop a tool to ‘examine motivation as a contextual influence on the ability of the intervention to improve health worker practice’ (p.3), it is by no means self-evident to me that a single composite index or score is (or can be) an appropriately valid and helpful tool to do so. In the domain of motivation, such a
single score may echo the attempts in the WHO’s WHR 2000 to construct a single indicator measuring a health system’s ‘performance’; for complex issues like motivation (like performance) such attempts are necessarily controversial.

Second, in the combination of constituting constructs (or latent factors), the question arises if these constituting elements can meaningfully be ‘added up’ to form a standardised ‘sum’ quantifying human beings’ motivation. In other words, there is a question of ‘(in)commensurability’. I am sure the authors are at least partly aware of this, as suggested by their own comment (p.9): ‘the SAQ responses in the 1st latent factor appear related more to workers’ commitment to the ideals of health work as a profession and less perhaps to actual, individual behaviours…’. If my interpretation is correct this relates to the SAQ statements like ‘I am a hard worker’; ‘I can be relied on by my colleagues at work’; ‘I am rarely absent from work’; ‘I do things that need doing without being asked or told’ (all of which get high scores – of course?). It is not that responses to such statements are irrelevant. My point would be that there is much more useful information in analysing the apparent contradictions (or lack of them) within the questionnaire’s set of responses, or in their comparison, than in the single sum of adding everything up and averaging. An additional argument supporting this last statement is, in my interpretation, provided by the mean 10 item scores in table 5. The overall average score for the 8 hospitals is 36.94 (on a maximum of 50) and the variation around this average is actually quite small. The highest motivation hospital’s score is 2.37 score-points higher; the lowest motivation hospital’s score is 1.03 score-points lower than this average. To me this means that the single score is probably not going to be very sensitive in capturing meaningful differences in motivation, be they across hospitals or over time. On the other hand, the differences between the scores in the different constructs are far more revealing, as shown by even a quick glance at Table 1. This might suggest that comparisons of the constructs’ or latent factors’ scores within the SAQ and across the 8 hospitals could provide much more useful information – even for the purpose stated on p.3 (‘to examine motivation as a contextual influence on the ability of the intervention to improve health worker practice…’).

Summing up these comments:

(1) The paper shows convincingly how a 23 questions SAQ can be made more parsimonious by reducing the number of questions to 10 and the number of groupings (constructs or latent factors) from 7 to 3.

(2) The paper is not convincing in showing that a summary of the questionnaire’s responses in the form of a single score is a useful or meaningful tool for capturing a composite measure of motivation. For this latter purpose the tool has not, to my understanding, been ‘validated’ (supposing this would be possible), and potential readers should be made clearly aware of this ‘deficiency’.

Clearly, a lot of careful data collection and thoughtful analysis has been done, and this can certainly be salvaged. A reworked version that would satisfy this reviewer would, however, involve relatively major changes.
Minor comments

p.5, para 3: “All qualitative data [27], [28]…”
The referencing looks awkward in this sentence; they do not add anything meaningful to the sentence

p.7, para 4: “We found a strong correlation of 0.9608”
If I read Table 4 correctly, this figure should be 0.9798

p.9, last sentence: “This clearly shows the value of using both subjective and objective measures of motivation for purposes of triangulation as well as validating the results acquired.”
If my understanding is correct, the use of the subjective and ‘objective’ measures of motivation, used as a triangulation device, in this case turns out to be invalidating rather than validating (related to the above ‘commensurability’ comment).

References (format)
I am not aware of any tradition in which there is a systematic inversion of the sequence name / given name between first author and subsequent authors (e.g. Martinez, J and T. Martineau) – but this may be my ignorance.

Table 1
Question 11: I would suggest to add an asterisk to the figure in the ‘mean score’ column (this question is also negatively worded)
Questions 15 and 16: the ‘mean scores’ are not well aligned with the ‘questions’

Table 2
The labelling of the hospitals (H4, H2, H1, etc.) looks garbled; there is no label H5, H6, H7 or H8, and H4 is used 4 times.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:
I declare that I have no competing interests